



10 May 1979

Wasted talent is wasted opportunity

THE Conservative and Unionist Party has been elected to govern the United Kingdom, with Mrs Thatcher as Prime Minister. The election was fought almost entirely on economic issues, and even on these the new government communicated little more than an intention. (We heard nothing about how much income tax would be cut, or what rise in value added tax would be necessary to pay for it, except for Labour's own, presumably biased, calculations.) Of science we have heard nothing—except on environmental issues from the Ecology Party, and on nuclear power from the Liberals.

We hope that the new government will recognise the central importance of science—and basic science, at that—to the economy. In opposition, the Conservatives have frequently called attention to 'wealth creation', which it says Labour has ignored. The Conservatives appear prepared to increase social inequalities (by redistributing the tax burden) and risk confrontation with organised labour (by introducing new legislation and cutting 'lame duck' industrial support grants). But are they prepared to invest in science? It is undeniable that science is increasingly the true source of wealth, now that the days of empire and exploitation are receding. President Carter has recognised this, with his constant calls to Congress for an increase in science funding. And West Germany has recognised it, with a massive increase in funding for its science ministry in this year's budget.

Shirley Williams, the outgoing Secretary of State for Education and Science, had also recognised it, managing to wrest a few extra millions for the research councils this year. But will the Tories recognise it? We hope they will have the vision to do so, particularly with Britain's first scientist (Mrs Thatcher was a chemist) as Prime Minister. Mrs Williams' balm merely alleviated the worst of a long history of financial stringency in the research councils, and there is a great deal of slack to be taken up—not least in the replacement of aging equipment like spectrometers and microscopes, or proper investment in Britain's great scientific successes like molecular biology (where with genetic engineering there is enormous economic potential) or even the less obviously economic radio and high energy astronomy. Great international centres of excellence have a galvanising effect on the science of a country that goes far beyond their particular discipline. Even a geophysicist, for example, feels pride that the double helix was discovered in this country.

There is no doubt that science in the UK is depressed. This is the time to revitalise it. The argument that basic science should be left for better times is false. The country that now invests deeply in science, despite the state of the economy, will be the one that emerges from the present world recession ready to take the lead.

Furthermore, it's not just a question of money. If wealth creation is a matter of science (and its application through

engineering), science is a matter of scientists. And Britain has never had a coherent policy for the growth and encouragement of its scientific community—just the usual British compromise of a decision here, a regulation there, a patchwork but no policy. The research councils have done their best to regulate and redirect studentships and fellowships; but they can only tinker with the problem. There is an enormous—perhaps 5,000-strong—and vital community of bright researchers locked into a cycle of frustrating and demoralising short-term contracts, produced by the unwritten policy of using an excessive number of research students as cheap labour and the static university system blocked with middle-aged tenured staff. Of course 'the best' always get jobs—that's only a matter of circular definition. The real question is of the wasted talent, the real talent, that does not find effective employment, either in the universities or in industry.

This problem is not Britain's alone. Even Germany suffers from it, despite its enlightened attitude to science. But the country that breaks through the logjam first will have an immense advantage. We challenge the new government to find effective, creative employment for Britain's post-doctoral fellows. A dozen new whizz-kid centres of advanced study, some applied, some pure? A proper career structure for non-teaching researchers at universities and medical schools would be a good start. The objective must be to create new research schools, either attached to universities as non-teaching institutes or completely independent. Research doesn't grow at random. It depends on potentially creative branches having the opportunity to develop at the right moment. The United States took its great lead in science when Europe rejected its intellectuals during the time of Hitler; they landed in America and research communities spread rapidly around them.

US pre-eminence in theoretical sub-nuclear physics, for example, can be traced largely back to Enrico Fermi's school in Chicago. And how did Britain invent its way through the second world war? It gave men like R. V. Jones responsibility the moment they could shoulder it.

We should take exactly the same attitude to our post-doctoral community. It must be given the opportunity to spawn its own schools of research, in an atmosphere where there is—if not total job security—at least hope, and growth, and vision. The age of this community—reaching into the mid thirties—is just right to bear fruit. (Studies by Harriet Zuckerman, recorded in her book *Scientific Elites* (Free Press, New York: 1977), show that the average age at which Nobel prize-winners did their prize-winning work was around 36. Those few who were under 30 were usually already members of flourishing research schools.)

So we say this to our new Prime Minister—whatever your successes or failures with industry and the more immediate problems of the nation: remember the scientists—and you will be remembered. □



US academy denies threshold for radiation damage

A COMMITTEE of the US National Academy of Sciences, after many months of heated debate, last week confirmed its support for the hypothesis that ionising radiation has an effect on the human body that remains directly proportional to the dose even at very low levels, and that there is therefore no threshold below which such radiation can be ignored.

This decision has been reached by the academy's Committee in the Biological Effects of Ionising Radiation—(BEIR) which had been asked by the Environmental Protection Agency to reconsider estimates of the hazards of ionising radiation first made in a report published by the committee in 1972.

Pointing out the absence of clear scientific evidence on the shape of the dose-response curve for low levels of radiation—a source of much recent controversy over, for example, the hazards faced by workers in nuclear shipyards and power stations—the committee says that the linear dose-response relationship “emerges by default as the simple model whose use appears to be the least objectionable”.

This conclusion, however, and the estimates of increased cancer incidence to which it leads, is sharply contested by a number of committee members, who state in a minority report signed by five of the 22 members that “far too little theoretical information exists to serve as a reliable guide for extrapolation.”

Speaking at a press conference held by the NAS to present the report, Dr Harald H. Rossi of the Columbia University College of Physicians and Surgeons in New York, one of the minority group, said he feared that the report would “contribute to excessive, and potentially detrimental, apprehension over radiation hazards.”

In reply, Dr Edward Radford, of the University of Pittsburgh's Graduate School of Public Health, chairman both of the BEIR committee and of its somatic effects subcommittee, issued a public challenge to any member of the minority group, or any other individual, “to meet me in public debate to discuss these issues openly and to see who comes nearest to the truth.”

The committee was asked by the EPA to look at the available data on both the genetic and somatic effects of ionising radiation. With regard to the first of these, the committee says that it came broadly to the same conclusions as that reached in the earlier report, namely that one rem of parental exposure throughout the adult population would lead to an increase of between five and 75 additional genetic disorders (over a cur-

rent figure of about 100,000) per million live-born offspring.

“The first BEIR report used mice data exclusively to estimate the incidence of chromosomal aberrations. We have had a limited amount of human and monkey data that has since become available, and have come up with figures that are not significantly different,” said Dr Dean Parker, chairman of the genetic effects subcommittee.

By far the largest section of the report—a single chapter taking up over 500 of the report's 850 pages—is devoted to assessing the somatic effects of high- and low-LET (linear energy transfer) radiation, the first including radiation such as that emitted by alpha particles, the second by X-rays and gamma-rays.

The report says that over the past few years, new studies have modified some of the earlier views about radiation induced cancer in humans. For example, solid tumours—in particular breast, thyroid and lung cancer—were of greater significance than leukaemia in assessing the risks of whole body exposure, and partly as a result of this, it was now felt that the total cancer risk was greater for women than for men. In addition, there was increasing evidence that age is a major factor in cancer risk related to radiation exposure.

Addressing the effects of low levels of radiation—of the order of tens or hundreds of millirads—to which the general public might be exposed in addition to the natural background, the committee says that at present there is no way of telling whether this exposure is detrimental, and that in any case the somatic effects would be masked by environmental or other factors.

Given the lack of data, it says that “for most radiation-induced cancers, the possibility of a linear no-threshold dose-effect relationship cannot be excluded”, adding that, for low doses, the linear hypothesis is “consistent with plausible carcinogenic mechanisms at the level of the single cell.”

However, it points out that the linear hypothesis probably overestimates the risk from low-LET radiation (such as X-rays and gamma-rays), but may underestimate the effect at low levels of high-LET radiation, such as alpha-particles.

Referring to those who receive occupational exposure of between 0.5 and 5 rems a year, the latter now being the recommended maximum in the US, the committee says that “a discernible carcinogenic effect could be manifest” at these levels of exposure. (The average

whole-body dose rate to the 30,000 people currently working in the US nuclear industry is between 0.6 and 0.8 rems a year).

Applying the linear hypothesis to a variety of available data—in particular that obtained from survivors of the atomic bombs dropped on Nagasaki and Hiroshima—the committee estimates a lifetime excess of cancer incidence caused by low-LET radiation of between 192 and 756 cases per rad per million males receiving a single exposure, and between 344 and 1,031 cases for females exposed under the same conditions.

The number of excess fatal cancers was estimated as between 70 and 353 per million per rad per year for single exposure, and between 68 and 293 for continuous cumulative exposure. In comparison, a report published by the United Nations Scientific Committee on the Effect of Atomic Radiation in 1977 estimated an excess of fatal cancers of 100 per rad per million for both categories (the higher BEIR estimates are largely the result of assuming a cancer risk that increases with age).

The legitimacy of these figures is questioned, however, by a number of members of the somatic effects subcommittee. Following what are described as “unresolvable differences”, five have signed a minority report stating that the risk estimated given in the report for specific organs—many of which were derived from the Japanese studies—were excessive, and that these figures “might be used to justify radiation exposure but not to indict it.”

Dr Rossi claimed that there was much radiobiological and epidemiological data indicating that the linear hypothesis led to substantial overestimates of radiation risk. “There is therefore a danger in a report which includes any evidence of an increase when I feel that the risks are nowhere as high as those in the report”, he said.

Dr Radford agreed with Dr Rossi that the general alarm about radiation dangers associated with nuclear power were not entirely justified. “It is a risk, but it is not the end of the world by any means. People do things to themselves that carry a much greater chance of leading to cancer than the effects of nuclear facilities,” he said.

However, he defended the committee's decision to support the linear hypothesis for low level radiation, and argued that the issues raised about the interpretation of scarce scientific data were sufficiently important for the scientific community at large to become

involved in the debate. "Ionising radiation is a prototype of other environmental agents which may be carcinogenic, even at low levels, so a resolution of this matter is therefore important" he said.

In a separate study published earlier in the week, another NAS study group looking at the risks associated with

nuclear power concluded that according to evidence in the scientific literature, exposure from routine operations of the nuclear industry increased the cancer risk of the most-exposed members of the public by 0.1%

If the nuclear industry was to expand along currently expected lines, there would be between 165 and 255

extra cancer deaths attributable to nuclear power operations up to the year 2000. The academy report says that production of an equivalent amount of electricity with coal-fired power stations would result in an estimated number of more than 3,000 associated deaths.

David Dickson

Weapons laboratories 'should retain links with university'

AN advisory committee last week recommended to the US Energy Secretary, James Schlesinger, that the University of California should continue to manage two laboratories which carry out the bulk of US research into nuclear weapons. However, the committee says that the university needs to improve the effectiveness of the "trusteeship" it has for the two laboratories, the Lawrence Livermore Laboratory outside San Francisco and the Los Alamos Scientific Laboratory near Santa Fe, New Mexico.

The committee has also advised the department to make a close study of alternative ways of managing the laboratories—in particular it suggests a private, non-profit corporation—if opposition within the university to the current arrangement should grow to a point which makes it "undesirable or impossible for the university to continue the present relationship".

The university's links with the two laboratories, whose activities grew out of wartime research but recently expanded into energy and environmental research, has long been a source of controversy. Faculty and students have attacked the arrangement, some on the basis that the university's responsibility for classified research is unethical, others objecting more specifically to the weapons research.

The current arrangement under which the university manages the two laboratories for the Department of Energy comes up for review next year. Last December, Mr Schlesinger asked a working group of the Energy Research Advisory Board (ERAB) to review the current relationship and recommend how it should evolve "to best serve the needs of the nation and the laboratories". In its report, which was forwarded to Mr Schlesinger by the board last week, the working group says that the factors which have made the university beneficial to the country and the laboratories are "lasting and fundamental", and that the relationship should therefore be continued and improved, even though the university needed to discharge more fully its trusteeship. Referring to what it calls a small but vocal opposition to continued nuclear weapons R & D in any form, it says that protesters often complicate the laboratories' already

difficult public relations by instigating press coverage of "peripheral issues". "The laboratories need support in this area which the university could help provide" the working group says.

However it adds that more pressures may develop both within and outside the university, making it impossible for the university to continue the present relationship. "The DOE must be prepared for such possibilities, even though we believe them to be remote," the working group says.

The report generated controversy when it was discussed at a meeting of the full advisory board in Washington last Thursday. Mr Tom Cochran of the Natural Resources

Defense Council, said he felt the working group had not dealt fully with the criticisms that had been made of the current arrangements, and that it had narrowed its focus by discussing merely the university's ability and apparent willingness to manage the laboratories, ignoring broader issues.

After considerable debate, the board agreed by 13 votes to 4 to forward the working group's report to Secretary Schlesinger, with a covering note from the chairman, Dr Saul Buchsbaum, explaining why the working group had interpreted its brief in the way that it had. A minority report is being prepared by dissenting members.

David Dickson

Anti-vivisection demo hits Cambridge

LAST weekend about three hundred people gathered in Cambridge—the place described by their leaders as the Mecca of biological research in the UK—to voice their disapproval of the use of animals in scientific experiments. For more than an hour they marched through the streets shouting their protests at empty laboratories and Saturday afternoon shoppers. They carried placards bearing slogans claiming "Cambridge scientists torture animals" and "your taxes pay for laboratory sadists to mutilate animals with no anaesthesia". The odd one called for the banning of the seal hunt and an end to factory farming.

The demonstration, which had attracted protestors from as far afield as Lancashire, Gloucestershire and London, was organised by Jean Pink of Animal Aid, an organisation concerned with the rights of animals, whose aim is to achieve the abolition of all animal experiments. "Vivisection is the ultimate in animal abuse", said Mrs Pink. "We demand the total abolition of the use of animals in laboratories".

The protest also included the delivery of a letter to the Vice-Chancellor of Cambridge University expressing the group's concern and a talk by Hans Ruesch from Switzerland, whose book *Slaughter of the Innocents* has recently been published. There is total ignorance about vivisection said Hans Ruesch. He claimed

Michael Alcott

Sorry, for copyright reasons some images on this page may not be available online

that few know of the "complete and utter uselessness of medical research through experimentation". Movements to seek the abolition of animal experiments had recently started up in several countries, he said, in particular West Germany and Italy, and also Australia.

The meeting ended with a request that the new Home Secretary, Mr William Whitelaw, who has just taken up his appointment after the election of a Conservative government, should be bombarded with letters as soon as he gets to his new office.

Judy Redfearn

Soviet 'mental illness' spreads to Czechoslovakia

Dr Snezhnevskii, of the Serbskii Institute of Forensic Psychiatry in Moscow, is well-known as the man who provided a theoretical basis for the practice of interning dissidents in mental hospitals. According to Dr Eva Dubska who left Czechoslovakia in 1977, at the time of her departure, Snezhnevskii's works were gaining support among certain sections of the Czechoslovak psychiatric establishment. A case history has now reached *Nature* which suggests that the Czechoslovak political establishment is prepared to add its weight to them.

According to Snezhnevskii, "all psychological illness is a form of schizophrenia. Further, he identifies a number of subtypes and symptoms not recognised abroad. These include "sluggish schizophrenia" (with no perceptible symptoms), "delusions of reformism", "a mania for protest" and the like. In the case of Augustin Navratil a 45-year-old Czech recently committed to the mental hospital in Kromeriz, there are disturbingly clear indications that the same diagnostic procedures have been used.

Navratil, who served as a local councillor for the revived People's Party under the Dubcek regime and who has since worked as a railway signal-box guard, was charged in January 1978 with slanderous the state. (He had been collecting signatures for one of the many civil rights petitions which sprang up in the wake of the Charter-77 movement.)

After considerable bureaucratic delays, he was sent to the local mental



"I'm afraid, doctor prosecutor, you are suffering from hysterical self-stylisation as a psychologist with an inability to adapt to an adequate view of humanity"

hospital for examination, and found unfit to stand trial; accordingly he was formally committed for treatment. The diagnosis includes such Snezhnevskii type expressions as "hysterical self-stylisation towards the ideal of a strong leading personality and with a strong moral responsibility which the subject understands as 'fidelity to his principles' and an inability to adapt to an adequate view of social reality". Further "by analogy with the fate of past personalities the subject thinks that 'for the truth one must logically suffer'."

At times the report becomes near ludicrous, quoting IQ figures as evidence of an unbalanced personality, urging treatment although admitting that Navratil cannot be brought to change his views. However, it is suggested, it may be possible to make him conform at least outwardly.

It is, perhaps, this Kafka-esque cynicism which is the most alarming aspect of the case. Reading the report, one gains the impression of a trained scientist desperately trying to adduce some kind of substantiation for a diagnosis already determined. □

French stress scientific side of detente in space

Mr Brezhnev's invitation to France to supply a crew-member for a joint manned space-mission with the Soviet Union did not come as a complete surprise to the Centre Nationale d'Etudes Spatiales (CNES). Shortly before President Giscard d'Estaing left for his "official working visit" to the Soviet Union, the chairman of CNES, Professor Hubert Curien, said that CNES was considering the possibility of Franco-Soviet cooperation in materials-processing in conditions of weightlessness.

CNES welcomed Mr Brezhnev's proposal in principle. They are not, however, interested in France simply being another country to have one of its citizens in orbit. "We are a scientific group", a CNES spokesman told *Nature*. "Such cooperation could be quite interesting, but we want to know what programme or programmes would be involved. We have a lot of interesting programmes in hand, and if the cooperation comes to pass, it would probably be one of the people in charge of one of these programmes who would take part."

No date has yet been even tentatively set for such a mission, nor have any details been worked out. But if CNES stand by their wish that "he would be an experimenter rather than a pilot", a significant change in Soviet procedure would be involved. Since the

three crew members of Soyuz-11 died tragically during re-entry, Soviet cosmonauts have always worn space-suits aboard the *Soyuz* transport vessels. The extra bulk means that only two cosmonauts can be launched in one craft. Although one is officially designated the pilot, the other is trained to be able to take over the controls in an emergency. CNES, however, would appear to prefer their representative to be essentially a passenger, ferried up to an orbital station where he would then proceed with his experiments.

Although CNES, as scientists, deprecate the publicity value of participation in the Soviet space programme, both the French and the Soviet governments seem to place great stress on it as what Mr Brezhnev last week called a "barometer of detente". France has occupied the leading position among non-Comecon countries in bilateral space cooperation with the Soviet Union.

The first general cooperation programme to include space was signed in 1966, and was presumably intended by the Soviet Union to encourage de Gaulle in his attempts to detach France from the US and NATO. Although French foreign policy has changed under succeeding presidents, the general world climate of detente and the impetus which the Franco-Soviet

space programme had already built up ensured its fruitful continuation.

France has, accordingly, taken part in a number of joint projects with the Soviet Union, including a laser-reflector experiment to determine accurately the distance to the Moon, a study of the atmosphere of Venus, investigation of solar radiation from Mars probes, studies of the magnetosphere, solar neutrons and gamma rays, and a programme of simultaneous rocket launches to study the *Aurora borealis* and *australis*.

At present, aboard Salyut-6, cosmonauts Vladimir Lyakhov and Valerii Ryumin are responsible for a France-Soviet materials-science experiment to synthesise magnetic alloys from materials which do not give rise to magnetic compounds in conditions of Earth gravity. This type of experiment will certainly be in the CNES planners' minds when they work out their programme for their "experimenter". Before he dons his space-suit, however, one important point remains to be settled by that branch of the Academie Francaise which regulates neologisms and loan-words. Is it the nationality of the crew-member or the registration of his craft which determines his designation? In other words, will he be officially described as an astronaut or a cosmonaut? □

news in brief

Carter urges expanded nuclear power programme: President Carter has told members of the US Congress that he strongly supports an expanded nuclear power programme, and that despite his continued opposition to the liquid metal fast breeder reactor at Clinch River, Tennessee, he is still in favour of the development of a more advanced fast breeder. "I want to emphasise that my opposition to the CRFBR does not imply my opposition to breeder reactors in general or to nuclear power", the President wrote in a letter to Congressional leaders asking them to support his attempts to halt the Clinch River project. "We need to pursue a vigorous programme of breeder reactor research and development so that this option can be commercially available to us if and when we need it." The President's remarks became publicly known two days before a mass anti-nuclear demonstration was held in Washington.

Official BNFL response on Windscale leak: Sir John Hill, chairman of British Nuclear Fuels Ltd, has responded to government enquiries about the highly active Windscale leak first reported on 18 March. In a letter to Tony Benn, outgoing Energy Secretary, Hill outlined the history and current status of the leak. High activity of 200 millicuries per litre was first recorded from borehole number 63 in November after four previous samplings done earlier in the year showed a lower activity of 0.2 mCi/l. This lower activity "did not cause particular concern" says Sir John because it was consistent with contamination due to a spill occurring twenty-five years ago. By March, investigations showed that the new activity was consistent with activity discovered in an out-of-use steel-lined sump located in an annexe where buffer storage tanks are housed. The route from the sump to the borehole is still unknown, however. The letter states that laboratory tests show that the "mean" migration rates of ^{137}Cs and ^{90}Sr in soil are about 1 metre in "several hundred days" so that activity will have decayed "to an insignificant level before the site perimeter is reached".

National Cancer Institute criticised over carcinogen testing: The US National Cancer Institute in Bethesda, Maryland, has been strongly criticised by the General Accounting Office, the investigating arm of the US Congress, for deficiencies in the managements of its programmes for testing substances for possible carcinogenicity. In a report published last week, the GAO says that in particular the Institute failed to spot serious deficiencies in testing programmes carried out on a contract basis by private research companies. The GAO said that many of these companies used unacceptable procedures to test chemicals, and that many test results had had to be thrown out by the Institute, at a cost of millions of dollars. Another report published by the GAO has criticised the lack of sufficient training in professional ethics given to employees of the National Science Foundation, particularly over activities relating to dealings with possible future employers. The GAO says that in discussions with programme managers in the NSF, investigators found that it was not unusual to be asked for an interview with an organisation for future employment while involved in reviewing grant applications or monitoring ongoing work for that organisation.

DHSS proposes increased job security for research staff: Professor A. J. Buller, Chief Scientist at the UK Department of Health and Social Security, has proposed a new scheme to save the jobs of scientists faced with unit closures. Under the scheme participating directors of DHSS-supported units will consider researchers who are made

redundant by unit closures for any vacancies that arise—before they invite applications from other candidates. Directors of units are under no obligation to hire anyone they consider to be unsuitable. In addition the DHSS will make extra posts available to directors pending vacancies in their normal complements. The DHSS will also pay for researchers' removal costs. But researchers will be expected to move to any part of the country and if they refuse, "even if there is only one offer they will be made redundant". The proposal has been circulated to the 34 DHSS-supported units and will be put into operation when 20 units have accepted.

Australian trade unions act to stop proliferation: Mr Cliff Dolan, vice-president of the Australian Council of Trade Unions, announced last week that three major trade unions would withdraw their labour to halt two new uranium projects in northern Australia. Citing the dangers of proliferation, waste disposal and health hazards, the ACTU has opposed the mining and export of Australia's considerable uranium reserves estimated to be 20% of the Western world's supply. The Amalgamated Metal Workers and Shipwrights Union, the Australian Railways Union and the Electrical Trades Union have agreed at recent meetings to follow the ACTU's recommendation. The unions will stop manufacture of mine equipment and its transport to mining sites.

Sweden and China to cooperate: The Royal Swedish Academy of Sciences and China's Academia Sinica have signed an agreement on cooperation in basic sciences. The agreement, signed in Peking during a recent visit by Professor Carl Gustaf Bernard, secretary of the Swedish Academy, comes into force on 1 July and provides for mutual exchanges of scholars. The Swedish Academy of Engineering Sciences has signed a similar agreement with the Chinese. In a discussion of Chinese science policy, Vice-President Fang Yi told Bernard that he was trying to restore the damage done to basic science by the Gang of Four. Training scientists in the West was a good way of tackling the problem. Several hundred scientists will probably visit Sweden under the new agreement, beginning with ten interested in methodology in basic research. Another area in which the Chinese want Swedish help is geology. Developing their natural resources in one way of paying for their modernisation programme. In response to a question on environmental pollution, the Vice-President replied that in a period of enormous expansion, it was more important to build up the industries than to confine their pollution.

India's new radio telescope operational soon:

By the end of this year India's new aperture synthesis radio telescope will be fully operational. It is the latest addition to the Tata Institute of Fundamental Research's (TIFR) Radio Astronomy Centre at Ooty. Like the previous Ooty Steerable Radio Telescope, which was commissioned in 1970, its design and construction is completely Indian. Its resolution will be equivalent to that of a 4 km diameter parabolic dish. The new telescope is an assembly of seven antennae which will be linked to the main Ooty Steerable Radio Telescope. The total estimated cost of the additional antennae and associated equipment is 4 million rupees (£180,000), an amount which makes it one of the cheapest aperture-synthesis radio telescopes.

From Dilip M. Salwi in New Delhi



Energy search comes down to earth

David Dickson reports on US hopes of extracting heat directly from the earth's crust

HOT dry rock, its supporters believe, offers one of the few practical, medium-term alternatives to fossil or nuclear sources of energy with little of the environmental problems or shortages associated with either. If it can be tapped efficiently, the potential energy available from the heat stored in the Earth's crust is enormous. It has been estimated that the energy supply for the US for a whole year is contained in a 50 cubic mile block of hot granite.

"We also think that it is the cleanest source of energy on the horizon," says Mr Greg Nunz, recently appointed manager of the Department of Energy's National Hot Dry Rock (HDR) Program at the Los Alamos Scientific Laboratory in northern New Mexico, where research into possible extraction techniques has been going on since the early 1970s.

At present, the most traditional and widely used method of extracting geothermal heat is through tapping underground reservoirs of hot water. The Los Alamos scientists, however, are developing ways of extracting heat from rock which is usually a result of radioactive decay, or of molten magma below the Earth's crust. The technique is to pump water down from the surface into a fracture that has been artificially created in impermeable rock. The water absorbs the heat from the rock, and is pumped back up to the surface again under sufficient pressure to prevent it from boiling.

"The concept is really very simple: we are literally trying to 'mine' the heat by creating a heat-exchanger underground", says Mr Nunz, pointing out that HDR techniques are therefore more flexible than those which require the existence of natural hydrothermal reservoirs.

Soon after discussions began at Los Alamos about the potential of HDR, field investigations were carried out into the high thermal gradient characteristics of the nearby Jemez Plateau, site of the third largest extinct volcano in the world. A preliminary test hole was drilled to a depth of 785 metres, where the temperature of the surrounding rock was found to be 100 °C. A second hole following two miles away at a site known as Fenton Hill. This found a temperature of 197 °C at a depth of 2,929 metres.



Steam is released at the hot dry rock facility, Los Alamos

The next stage was to pump water under pressure down the drill-hole, opening up plate-shaped cavities—the largest having a radius of 215 metres—at the bottom. A further hole was then drilled from the same site in the hope that it would intersect the largest cavity and provide a channel for hot water to be returned to the surface. Uncertainty about the exact shape of the cavity, however, meant that the new hole failed to make the intersection. Mr Nunz explains that "the instrumentation of the time was not up to the job".

The next two years were spent developing techniques to get a better understanding of the geometry of well-bores and of fractures. In early 1977, the first of the Fenton Hill holes was re-drilled, this time successfully intersecting a fracture made at the base of the second. And late in the year the first circulation of water took place, using a heat-exchanger at the surface to disperse the heat brought up.

Tests carried out so far on the completed system have been a "tremendous success", says Mr Nunz. For example, during an initial 20-hour run the temperature of the water reaching the surface rose by 130 °C. Water quality was good, with a low concentration of minerals that could have clogged up the system.

This was followed by a 75-day test at the beginning of last year, during which the amount of energy extracted rose to about 5000 kW. The amount of water lost to the surrounding rock stabilised at no more than 1.5% of the circulation rate.

The success of this trial system led the Department of Energy to announce last autumn that it was setting up a national Hot Dry Rock Geothermal Energy Program, to be managed by the Los Alamos Scientific Laboratory—itsself run by the University of California for the department—with the overall aim of determining the potential of HDR as a significant energy source, and if possible of providing a basis for its commercial development.

The scientists at Los Alamos are already convinced that energy produced in this way—at least that which

can be used for heating—is potentially competitive in price with other more conventional sources of energy, and is likely to become increasingly attractive as the price of the other sources continues to rise. A recent study carried out with economists from the University of New Mexico, for example, reached the conclusion that in areas where the geothermal gradient is 40 °C or more per kilometre depth, electricity produced by HDR techniques would be competitive with other sources producing energy at 3 cents a kilowatt-hour.

The scientists also claim, that, so far, they have not been able to detect a significant environmental impact from the work on the Fenton Hill site. There has been virtually no disruption of local ecosystems, neither has the drilling caused any measurable seismic activity.

The next step in the national programme is to see if the results achieved at Los Alamos can be repeated elsewhere. Already three of the 100-mile square sites in different parts of the US are being examined as possible candidates for a second drilling programme. Meanwhile back at Fenton Hill, work has already started on a deeper hole, drilling to temperatures around 260 °C, which, it is hoped, will be able to demonstrate the commercial potential of HDR to interested utility companies. If successful, this could lead to a 50 kW experimental plant, sufficient to provide the electricity needs of 5,000 people.

One problem that the programme managers face, however, is gaining public recognition of the potential of HDR. "Right now, public perception of the potential is almost zero. And we are planning to have a contractor look at the problem and decide the most cost-effective way to do some public education," says Mr Nunz.

Among those to whom the message will be directed are likely to be members of Congress. So far federal support for the HDR programme has been extremely modest compared to the amount spent on research on other energy sources. Funds for the first few years' research came from the Laboratory's discretionary money, and

over a seven-year period the whole programme has cost under \$40 million.

Already other countries, particularly those with favourably geological conditions, have shown close interest in the programme. Once a year the programme managers hold a formal seminar in Sante Fe to discuss the progress of the project, attended by scien-

tists from many parts of the world. Both the Japanese and West German governments have offered to provide funding for the programme.

The federal programme is now focused on 1986, when a decision will be taken on whether to move the whole programme up to a near commercial scale.

At Los Alamos, they are confident that these investigations will be successful. "I don't think that the world has discovered hot dry rocks yet. But we see it as the only major alternate energy source that can be developed in this century, and do not understand why the powers that be are not more excited about it," says Mr Nunz. □

Harrisburg shakes up Sweden's nuclear war of words

THE Harrisburg reactor accident hit the Swedish nuclear energy debate like a karate chop. The Social Democrats, long supporters of Sweden's nuclear development and opposed to a referendum on it, announced that the accident was serious enough to make them change their basic judgements: there should be a Swedish committee of enquiry into Harrisburg and a referendum should be held. The Liberal party government changed with them and declared that no more reactors should be loaded until after the referendum. And an opinion poll taken after the accident shows a dramatic anti-nuclear swing: 53% of people asked said that they would vote against nuclear power in a referendum and only 26% said they would vote for. In January, the same question produced 43% against and 41% for. On the face of it, then, Sweden's nuclear future is wide open again. Or is it?

The seven years of Sweden's public nuclear debate has shown that facts enter politics only to the degree convenient for either side. More facts are being gathered now: a party of Swedes and Danes left this week for Harrisburg, and, on the domestic front, a working group is about to investigate Good Friday's accident at the Barsebäck 1 plant, within sight of Copenhagen. In this accident—"the worst in Sweden so far", according to a spokesman at the Nuclear Power Inspectorate—the generator exploded, throwing out metal projectiles (including a two-ton steel capsule) which did not, however, penetrate into the reactor itself. But it is hardly likely that any of the facts uncovered by these groups will be treated any differently from those already paraded. The referendum, to be held next spring, will be advisory—but its 'advice' will no doubt be used only so far as the politicians find convenient.

The last act of last year's three-party coalition government was to refuse the nuclear industry's application for the loading of the Ringhals 3 and Forsmark 1 reactors, on the grounds that the industry had not shown the existence of a sufficiently large rock formation with appropriate characteristics for final disposal of waste. The government said that, if the industry found rock they

considered more suitable, and if it is approved by the Nuclear Power Inspectorate, the government would give permission to load the reactors. After the coalition fell, and the pro-nuclear Liberal party government took over, the industry duly made more inspections on Sternö, near Karlshamn in southern Sweden, and made a new application.

The Inspectorate decided to appoint a group of eight independent geologists (six from Sweden, the others from Norway and Finland) to assess the industry's report, giving the Inspectorate's Board a factual base for its recommendation to the government. The new government announced that it would be bound by the Inspectorate's recommendation—and that it was confident that the Inspectorate would be able to recommend that the loading should go ahead.

Rock unsuitable

Seven of the eight geologists found that the geological properties of the Sternö rock were such that it "cannot be used for the storage proposed by the nuclear industry". They reported: "The holes drilled (by the industry) are too few and drilled in such directions that, even if they had all shown fracture-free, homogenous rock, they would not have sufficed to show unambiguously that the rock fulfils the set conditions." The group found fractures not discovered by the industry, and these could reduce the area usable for storage from the desired 1 km² to 0.3 km². Although storage requires homogenous rock, the investigated area contained three different types of rock, and the consequences of this had not been sufficiently investigated. The group maintained that it is impossible, at present, to determine the permeability of the rock—yet the nuclear industry's safety analysis depends on this parameter.

But the Inspectorate's Board did not recommend that the loading be stopped. In a masterly formulation, it told the government that the geological qualities of the rock must be seen in the context of *all* the barriers designed to prevent leakage of waste including vitrification and encapsulation. "The

Inspectorate does not maintain that each one of these barriers must give that degree of protection assumed by the nuclear industry in order for the final result to be regarded as wholly secure. . . . The importance of the demands made on the geological barrier should not be exaggerated . . . if the other barriers work satisfactorily." The question of whether the industry had shown that the rock met geological standards for deposition of nuclear waste was neatly sidestepped. The Board said that it was not unanimous on the meaning of the word "show", and therefore could not say whether the standards had been met or not.

"This has been such a political affair that science has been totally over-ridden", says Dr Arne Wesslén, one of the eight geologists. "It is regrettable that a regulatory authority can appoint a group of experts who give such a consistent opinion, and then ignore it for political reasons." Dr Bengt Åberg, another of the geologists, agrees. "We did the work scientifically. The Inspectorate tried to find ways around certain parts of our report", he says. And a spokesman at the Inspectorate admits, "We found ourselves under intense political pressure".

Ironically, all the geologists think that it should be possible to find rock in Sweden with the desired characteristics. "Sternö was not chosen for geological reasons", says Dr Åberg. "The industry has had to rely on areas it owned or where it was given permission to drill. On Sternö, no holes could be drilled outside the power station fence! We should have a law allowing drilling to be done in geologically-likely areas."

The day after the Inspectorate's Board met, the accident happened at Harrisburg. In the aftermath of that, with the ban on loading more reactors until after the referendum, the Board's verbal acrobatics proved to have been unnecessary—for the time being. Politicians' energies are now concentrated on the fight over the wording of the referendum question, and bitter fights lie ahead. In the meantime, as September's general election looms, nuclear power is not likely to be forgotten.

Wendy Barnaby

UNCSTD '79: This is the first of a series of articles leading up to the opening of the United Nations Conference on Science and Technology for Development (UNCSTD) in Vienna on 20 August. So much hot air has risen over the nature of the conference itself that the real problems facing science and science planning in the Third World have been forgotten. So we sent a team of writers to the grass roots. We asked Dr Hanlon

to look at the problems, on the ground, of an African country. He chose Ghana. We asked Anil Agarwal, a respected Indian science journalist now working in London, to visit India to report on the changing shape of science under the Janata government. And we have granted writing fellowships to scientists to visit and report on affairs in South America. We shall also have reports on South-East Asia and the Middle East.

Ghana: struggling for scientific independence

"TWELVE years ago I came back with a PhD in magnetohydrodynamics. I tried to work here, but there was no point. It was totally useless in Ghana." Now K. O. Kessey is professor of mechanical engineering at the University of Science and Technology (UST), working in refrigeration. "Twelve years ago, I couldn't imagine me doing earthly things—I was highly mathematically inclined. But you must use your knowledge and experience in a new situation. Students come back from abroad and they are helpless and hopeless. They say 'Oh, I don't have the equipment to continue my work'. So they do nothing. But it's just an excuse."

Dr Dwuma-Badu, head of the UST pharmaceutical chemistry department, and one of many working on herbal medicine, said: "With the equipment here, we cannot challenge pharmacy schools in developed countries with synthesising new compounds. So we work with Mother Nature to find what is there."

These comments reflect a growing change in Ghana. Science there used to be closely modelled on the British pattern; but as more graduates returned from studying abroad they began to question the relevance of their training—and of Western science in general. Meanwhile, the government, which had previously ignored science, began to ask what it was getting for its money. Both sides began to realise that Western science had simply been transplanted and allowed to grow in isolation, producing neither good science nor anything useful for the country.

As a result, the government set up a committee to review the Council for Scientific and Industrial Research (CSIR), and after its highly critical report, CSIR sharply changed direction. The Five Year Plan published two years ago was also the first Ghanaian development plan to have a section on science and technology.

As Dr M. N. B. Ayiku, head of CSIR's Industrial Research Institute explains, "there is only one science—only one set of laws of nature. But in a particular society, a certain type of science is emphasised. There is bound to be an accent based on our own problems. We must start small and

First World universities still set the pace for science in Africa.

But the resulting research is often irrelevant

Report by Joseph Hanlon

solve the little problems first. But as you solve the little ones, you unearth the big ones." And that brings new areas of basic science into view.

Professor Albert N. Tackie, head of CSIR, points to the other side: no one else will solve Ghana's problems—such as malaria. "If we want to tackle our diseases, we will have to do the research." Similarly, he points to building, where Ghana still relies on imported cement when it should be concentrating on local bricks and timber. The latter requires the study of another problem that is much less serious in temperate zones and thus little researched—termites and other insect pests.

Basic research—now underway on several Ghanaian problems, may have application elsewhere, for example:

- At UST, the chemistry and biology departments are working on bacteria to leach gold ore from the sulphur in which it is embedded.

- At the University of Ghana geologists are studying an earthquake fault which runs through Accra. It has not been active until recently, and its activity is not explained by current theories, so it is of general geological interest. The research also has practical value for Accra, where 16 people were killed in the 1939 earthquake.

- The Food Research Institute is looking into the preservation of kenky, fermenting corn meal, which must be sold and eaten in just a few days before it becomes too sour. The Institute is experimenting with dehydration to improve storage and distribution. Little work has been done elsewhere on drying fermented doughs, because they are rarely eaten in developed countries.

- Ghana is well-advanced in the study of herbal medicines (the subject of a later article).

Not all of Ghana's scientists have made the transition; it has been hard even for those who support the changed policy. J. Maud Kordylas, head of the Food Research Institute,

talked of her own research. She is now working on the use of the winged bean as a weaning food. Nutrition studies have shown a definite need for a supplemental weaning food, but peasants cannot afford to buy commercial ones. The winged bean can be grown locally and is an important source of the amino acid methionine. Kordylas's work on processing the bean is very different from what she was doing before. "I worked on vitamin A metabolism with respect to cholesterol at the London School of Hygiene and Tropical Medicine. I see high blood pressure in Ghanaians and I know they eat a lot of palm oil which is high in vitamin A, so I know it would be useful. And that sort of work gives me more satisfaction as a person. But we have many nutritional problems. In practice, the winged bean is a more useful project for Ghana."

Postgraduate studies abroad remain a necessity, but Ghana now scrutinizes these more closely. CSIR sends many of its research officers abroad for MScs and PhDs, and four years ago, it began vetting their research projects to make sure they are relevant, according to Tackie. Nicholas Darkwa, for example, has recently returned from North Carolina University with a PhD in pulp and paper science. His research topic was making paper from the stalks of plantain, a banana-like plant very common in West Africa, and CSIR actually shipped plantain stalks to him in the US for his work. Another CSIR researcher is studying making paper from local hardwood trees as part of his work in the USSR. Tackie hopes that the two will form the scientific basis of a Ghanaian paper industry.

The other side of education abroad is the brain drain, both short and long term. Foreign countries may subsidise Third World students, but they extract a high price—encouraging the brightest to remain abroad. UST vice-chancellor E. Bamfo Kwakyie comments "to send your best students out means that you neglect your own postgraduate activities. Furthermore, it costs 24,000 cedis a year to keep a postgraduate student in the US—we pay a professor less than half that. Finally, we have severe staff shortages." So UST has changed its policy: fewer students are sent abroad, and an effort is being

made for joint MSc programmes, in which the student returns to UST to finish his or her research.

Although the new policies of UST, CSIR, and others are working well in some areas, in others they are not. UST has a joint chemistry MSc programme with a foreign university which simply allowed UST students to enrol there for a PhD rather than return to Ghana, defeating the whole point of the programme. A physicist told me he was working on electron transport in thin manganese films, a project of no apparent relevance to Ghana, because it was the area of a visiting British professor. Manganese may not be useful for Ghana, "but for someone who wants to get his PhD and get promoted, it is useful," commented the physicist honestly. Promotion is always a bugbear in science, and it is a special problem in Ghana, both because the system is very rigid and because it has failed to take into account the new attitudes toward research. "People are employed on the basis of being world class scientists who are employable in Britain or the US, not on whether they can solve local problems," declared M. D. Swain, a Ghana University botanist.

Promotion at CSIR, for example, is based on four things: a list of publica-

tions, reports on those publications by outside assessors who are primarily foreign, annual assessments, and supervisor's comments. There is little consideration of work that may have been highly useful but has not led to publications. Ayiku complained: "If I want promotion in CSIR, my *curriculum vitae* is sent to Germany or England, where I am judged by European standards." CSIR claims to have changed its policy somewhat, using a few local assessors for papers, for example. But few people, inside or outside CSIR, seem aware of the change.

In practice, the universities are just as bad. "It shouldn't be the *sine qua non* for promotion to be published in prestige Western journals," admits Kwakye—but it still is.

Often, because of rigidity, places simply do not get filled. Several UST departments do not have professors, even though they are being run perfectly well by acting heads who only lack the paper qualifications. One University of Ghana department has four of its 11 senior positions vacant. Yet, when a highly competent man applied, he was rejected—because he had failed his first year of university, even though he repeated it successfully and has since done well.

There are few local journals and they are rarely published because of lack of foreign exchange, so publication means publication in a foreign journal. One botanist considered looking at sour grass, a weed first noticed in Africa only two years ago and now slowly spreading out from Accra. It would have made a useful piece of scientific work, ideal for a local journal; but lacking international interest, it would not qualify for a foreign journal. So the work was not done.

Another adverse factor is the lack of an outlet for the results of relevant research. Scientists complain that, despite the lip service paid in the Five Year Plan, research is still not taken seriously. Ghana's national paper for UNCSTD, written by Tackie, complains that "the major scientific and technological institutions such as CSIR and the National Standards Board are not involved in investment decisions . . . development planning is entrusted almost exclusively to economists. . . . There is no effective accompanying technological evaluation."

There is little industrial take up of ideas generated by researchers, either. UST set up a few small manufacturing units on campus using simple equipment and local materials. The engineering faculty makes pumps and traffic lights and the arts faculty has a ceramics production unit. "Originally, we hoped that entrepreneurs would take these out of our hands, but this has not happened, so we continue to manu-

facture," Kwakye said.

Continued dominance of the economy by foreign companies also reduces the opportunity for local researchers. "The multinationals are so powerful; Ghana is so poor. It is impossible to think how we shall break this hold," comments Kwakye. Much applied and basic research is done outside the country or by outsiders. For example, when off-shore oil explorations began, the international oil companies asked local geologists to give them a basic briefing, and then they brought in their own geologists. So the local geologists gained nothing out of the project except a few small fees.

The scientists themselves are partly to blame. The ecology and geology of Ghana remain poorly studied. Tropical forests are Ghana's second highest foreign exchange earner, yet they are little researched. Tropical ecology is different and much more complex than temperate, and is sure to become a prime research area as the temperate zones are better understood—indeed, much basic science in the future will probably have a tropical bias. But it is not the Ghanaians who are moving into these new areas, but the white expatriots and visitors. Even the research into earthquakes is dominated by people and equipment from London University.

Some change does seem to be taking place in Ghana, however. There has been a realisation, as the UNCSTD paper notes, that "some of the measures taken so far towards the goal of economic independence have paradoxically led to a greater degree of dependence on the industrialised nations." Ayiku notes that Ghana has been buying technology as if it were Coca Cola—to be consumed and thrown away with no understanding of it. As Ghana University geologist Mike Mensah points out: "a developing country that cannot do research will always be an importer of technology."

But, as the UNCSTD report also notes: "The Ghanaian experience suggests that mere growth in the volume of scientific research or the use of advanced sophisticated industrial plants or machinery will not of themselves bring about development. It is essential that scientific research should reflect the true needs of development."

This must be done with the facilities available in Ghana. Ivan Addae-Mensah, a University of Ghana chemist, concludes: "If you want to do high-falutin sophisticated research, you can't get anything done here. But if you actually want to be a scientist here, you must look for the problems around you." □

Dr Hanlon is a science journalist specialising in Third World affairs.



Above: Professor Tackie, research council chairman. Below: Professor Kessey with down-to-earth invention, a special carrier for keeping vaccines cold.



Why eating should carry a government health warning

It is time for scientists to swallow their doubts and press for a proper food policy which will save lives, say **John Rivers** and **Philip Payne** of the Department of Human Nutrition, London School of Hygiene and Tropical Medicine

It is generally agreed that food can kill, but argument persists as to whether anything should be done about it. Government action on diet and health has so far been minimal essentially because argument over the exact risks of eating has dwarfed any attempt at exploring the problems of doing something about them. We believe, however, that the latter issue is not merely as important as the former, but must be resolved first: we must know what kinds of action we might conceivably agree to before it is even worth considering the role of diet in disease.

It is widely agreed, by both nutritionists and the general public, that changes in diet occurring in the past 100 years have at least some role in the increased incidence of a whole range of diseases from dental caries and obesity to cardiovascular disease and cancer. These dietary changes have been many and complex, and major controversies exist as to exactly which are to be held responsible for the increase in diseases of affluence.

Considerable publicity has been given to such controversies, not least by interested organisations in the food industry who feel their own products might be under attack. Understandably such publicity as a rule is aimed at defending the notion that all existing consumption patterns are expressions of free choice, and should be regarded as innocent until proved guilty. This philosophy of waiting for absolute scientific proof attracts much support from the health professions, not least because it implies both the need for much continuing scientific research, and that the professions involved will define and apply their own criteria for what constitutes adequate proof.

Paradoxically the adherence of the health professions to a philosophy of waiting for certainty coexists with a broad professional consensus that would support moderation in eating, a reduction in fat and sugar intakes, and, in consonance with many other fashionable theories about conservation and ecological balance, some degree of reversion to diets with greater proportions of plant foods and fewer refined components. This consensus is clearly only provisional. That is, whilst there is not wide agreement that these changes will ultimately be shown as

sufficient, or indeed necessary to bring about a general improvement in nutrition, they are currently widely accepted as reasonable goals for any individual interested in conserving health. It is, in short, the best advice we can offer at present.

Unfortunately the nutrition profession in the UK has a somewhat vacillating attitude about what to do with its consensual knowledge. There is much pride about the role that nutritionists apparently played in the institution of a nutritionally based food policy during the second world war. There is widespread involvement with organisations like FAO and WHO in attempts to ensure that nutrition is integrated into the development policies of the Third World.

Yet concern for the nutritional problems of our affluent society is not translated into such decisive action. Admittedly individual nutritionists try to persuade, cajole and educate the public, but the profession has, as yet, stopped short of concerted action. It does so on the ground that present knowledge is inadequate. Clearly it is, but insofar as dietary moderation would tend to halt some of the major current trends of dietary change, it represents a policy of conservative caution, an attempt to stand still until the implications of change are better understood. This is a powerful argument to weigh against the elusive ideal of certainty.

It is generally agreed that piecemeal nutrition education usually fails, and some government action is needed. If the government were persuaded to support a strategy for health it would require at least consistent and relevant action by various ministries, and possibly regulations for the control of advertising. Such a sweeping approach might well be effective in changing the overall pattern of consumption, but is resisted on two grounds: it would constitute an unjustified interference with free choice; and the economic costs to agriculture and industry of adapting to new patterns of demand would be unjustifiable in view of the uncertain benefits. We reject both these objections.

In the UK, government intervenes in agriculture in many different ways, and any idea that we have a free market economy in food is a myth. The choice



Choosing can be difficult . . .

is between agricultural policies that ignore the health of the consumer, and those in which health is one of the considerations used to frame policy. At the moment, health enters only peripherally into, for example, regulations to prevent the adulteration of milk. The central complex of subsidies and support systems that shape agricultural policy in the UK seem to exist for wholly non-nutritional reasons. The subsidy on butter may make political sense, but, in a country where the consensus advice is to reduce dairy fat consumption, does it make nutritional sense?

The incentives for producing fat animals result in masses of fatty carcass waste that is used in the manufacture of cheap convenience foods such as sausages and beefburgers. A health conscious consumer has a difficult balance between cost, time and risk to cope with, and he may feel that choosing a diet he regards as unhealthy is an unavoidable risk. The refusal of the Milk Marketing Board to make reduced fat milks freely available; the absence of regulations requiring the labelling of cheeses with fat value; the widespread use of sugar in processed foods; the high price of wholemeal bread—all these results of government food and agricultural policies, or the marketing practices of virtually uncontrolled oligopolies, seriously impede any individual seeking to follow the sort of dietary advice which most nutritionists, doctors and the UK Department of Health and Social Security offer.

The existence of non-nutritional agricultural policies makes nonsense of the argument that inaction on diet and health should persist until final proof is



... especially when agricultural policies favour convenience foods

obtained. There is no doubt that the current nutritional debate is tentative and incomplete, but if nutritional factors are to be used with political and economic ones to frame agricultural policies, all that need be asked is that the predictions in each area should be equally plausible. To accept political intuition about the benefits of an agricultural policy while making a philosophically impossible absolute proof the criterion for scientific inputs is to emasculate scientific advice to government.

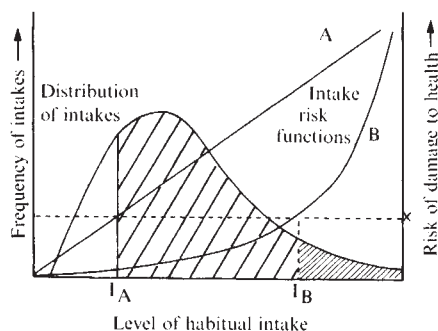
We believe that the health professions should unite in urging now a cautious and conservative basis for a nutritionally oriented food policy. Once they do so, however, there are other challenges to be faced, and it is important to pursue further some of the possibilities which expanding scientific knowledge will raise. It is reasonable to expect a hardening of the consensus view about the adverse health effects of some particular food or nutrient, and questions will then inevitably arise about further legislation, more specific advice or more rigorous action. If we consider now what kinds of scientific evidence would be needed to justify such intervention, nutrition research could be planned so as to contribute most effectively to the social and political processes of policy choice, and hence to health.

Clearly the risks of consumption of a given dietary component vary with the amount consumed. Since there is a wide variation between individuals in the amount of any dietary component that they consume, there will be wide variation in the risk to which they are

exposed. This will be increased further by variations between individuals in their constitutional sensitivity to disease. The benefits derived from a given 'across the board' reduction in intake will also, therefore, be highly variable.

If food and agricultural policies designed to combat disease are to be both effective and equitable it is necessary to define both the intake-risk relationship, and the acceptable level of risk in quantitative terms before it is possible to consider precise costs and benefits of the policy options open. The problem of what constitutes an acceptable level of risk is encountered in legislating against all sorts of hazards and the defining of acceptable levels is not a scientific problem, but a social and political one. However, the problem of intake-risk relationships does require scientific definition, since their nature can drastically modify the health policy chosen.

If the intake-risk function is linear (curve A in the diagram), everyone with an intake above I_A is above the acceptable level of risk, and in the example this constitutes the majority



of the population. If however the asymptotic function, curve B, applies only the small fraction of the population consuming more than I_B is at risk.

Clearly the difference in size of these risk groups means there will be great differences in the effectiveness of nutrition policies aimed at the whole population. Moreover, since measures to reduce disease will, in general, impose on those not at risk, the political implications, and the long term acceptability, of such policies will also differ markedly.

If nutrition theories are to be translated into nutrition policy, the quantification of the currently accepted relationships between risk factors and disease is more important than simply trying to increase our certainty about the strength of that relationship.

A change in the direction of nutritional research from causality to risk is a major one, and many scientists may view it with suspicion. But since in the foreseeable future nutritional knowledge will be chiefly applied through social policy, it is a change that must occur.

One further change seems probable. At present where nutrition policies exist at all, they have been evolved by senior administrators in closed counsel with leaders of the health professions. Such cabals will not survive for long. Misgiving about the role of scientists in government have grown to a general lack of confidence about "establishment science" as the debates over the Dangerous Pathogens Advisory Group or Windscale have made clear. It seems likely that general public concern about the role of nutrition in disease, will soon be matched by public demands to be involved in framing the policies necessary for change.

Such a change will underscore the irrelevance of the nutritionists' professional standards of certainty. Indeed they will strip away much of the veneer of professionalism that the subject has acquired. But once nutritionists recognise the essentially political nature of decisions about food policy, and the necessarily tentative nature of the theories that underpin them, we believe that they will welcome the involvement of consumer organisations and other pressure groups in the consultative process. It is reasonable that they should. At present while some scientific advisers strive after absolute proof in their decision making, the food industry employs others as spokesmen in the scientific debate, exerts massive influence on the direction of research by the funds it deploys, and by advertising pressures continues to manipulate food habits in ways that are widely suspected as having led to increased disease. A change in the way we tackle these problems is long overdue. □

correspondence

Select Committee's future

SIR,—In its recent First Report (17 July, 1978) the Select Committee on Procedure proposed some important and far-reaching changes in the procedure of the UK House of Commons. These include the establishment of a separate select committee for each of the departments of government, or groups of departments.

Our concern in this letter is with only one of the proposals, namely the dissolution of the Select Committee on Science and Technology whose work, it is said (paragraph 5.40), "should in future be carried out by the departmentally-related committees . . ." This proposal is supported in the report by an argument to the effect that the proposed new committees "would in no way be inhibited from taking evidence from any government department or other body relevant to the studies which they undertake."

Nevertheless, it is the view of the Executive Committee of this council that such a decision would be a retrograde development. We strongly support the opinion expressed to the Select Committee on Procedure by Mr Arthur Palmer MP, the Chairman of the Select Committee on Science and Technology, as indicated in paragraph 5.40 of the First Report. Departmental committees, said Mr Palmer, could not "adequately meet the need for parliamentary scrutiny of scientific and technological developments" since "inter-departmental subjects, however important they might be, would be accorded a lower priority".

The Committee on Procedure's argument depends on the generality of the policy of spreading science research and development amongst the departments of government. From many points of view there are good reasons for continuing this policy, where "after brief experiments in the 1960s with co-ordinating ministries, governments of all complexions have decided to leave scientific and technical research in the hands of individual ministries with only vestigial co-ordinating functions remaining at the centre" (paragraph 5.19). It does not follow from this, however, that the *scrutiny* of these activities should be compartmentalised to the same degree. Indeed, it is precisely because the central coordinating functions for scientific and technical research are vestigial that the work of the Select Committee on Science and Technology is so valuable.

Presumably, within the proposed system, there would be parliamentary committees for each of the Departments of Education and Science, Energy, Environment, and Industry, and perhaps other departments as well, all dealing with particular aspects of science and technology. Such committees would, undoubtedly, serve a valuable purpose in scrutinising the day-to-day, ongoing, short-term, research that is central to the work of each department. But no amount of "liaison and cooperation" between such committees "to allow joint inquiries and consultation, where committees share an interest" (paragraph 5.20), could deal satisfactorily with the broader, long-term, issues of

science and technology.

This can be seen at once from the issues that have already been taken up by the Select Committee on Science and Technology. These were necessarily concerned with scientific and technological questions drawn from a whole range of departments such as energy, environment, industry, health and social security, etc. Current issues of great political, social and economic significance, such as the potential benefits and hazards of genetic manipulation or the future effects of electronic microprocessors, obviously overflow all the boundaries appropriate for the administration of science and research within the government machine. By its terms of reference, which encourage it to scrutinise positively the activities of the executive on such issues right across the board, the Select Committee on Science and Technology fulfills an indispensable role on behalf of the lay public which its members are elected to represent. From the experience of our council and the detailed studies that we have made of policies and procedures relating to the social and political context of science and technology, it is clear that this is a role that should be considerably strengthened, rather than weakened, as proposed. As an institution endowed with all the constitutional authority of Parliament, capable of mediating effectively between the various domains of scientific expertise, governmental administration, and the concerns of the general public, the Select Committee on Science and Technology is potentially one of the most valuable social instruments available to us in times of rapid and bewildering technical change.

In the past the Select Committee on Science and Technology has not always been able to achieve this full potential. From our own close contact with members of that committee, and from our observation of its work, this is not due to any fundamental defect in its terms of reference, but rather to the very limited permanent staff resources, and restrictions on the appointment of specialist advisors, under which the committee has laboured. These deficiencies are recognised in the First Report from the Select Committee on Procedure, which recommends much improved arrangements in this respect. In our view, it is ironical that these resources should now be made available to new select committees which are to be much more closely linked to particular departments where technical staff resources abound, whilst proposing to terminate a committee which has been weakened by lack of such resources.

In our view, the proposed new system would perpetuate, in a most unsatisfactory manner, the split between science and technology which occurred when the DES was first created. Furthermore, there are many areas within technology itself which cannot be neatly divided into the interests of Departments of Energy, of the Environment, of Industry, etc. This fragmentation is not merely inconsistent with the realities of modern scientific research and technological innovation. It also takes a narrowly instrumental view of the contributions of various expert

disciplines to the issues raised by technical change. It is not so much, as the report suggests (paragraph 5.42), "that scientific and technical matters in widely varying fields involve a common approach which can best be handled by common methods"; it is rather that many current issues and policies cannot be adequately scrutinised except from a point of view that transcends and integrates the separate departments of government. The various proposed departmental committees, however loose their terms of reference and however well "coordinated" they might be over particular issues, could not adopt a sufficiently comprehensive viewpoint where social, environmental, economic, health and safety and industrial factors would be related to narrower technical or administrative matters.

We hope, therefore that Parliament will consider very carefully before dissolving the Select Committee on Science and Technology. Whether or not departmental committees are created, we believe that the Select Committee on Science and Technology should be retained. Its broad remit it seems to us, is just what is needed for the purpose of taking an overview of the various areas where pure science interacts with all the separate branches of technology, and where these branches interact with each other and where they may also have far-reaching social consequences.

JOHN ZIMAN KENNETH DENBIGH
Council for Science and Society,
London, UK.

Scientist-entrepreneurs

SIR,—One aspect was not mentioned in the article by Dickson (5 April, page 494) concerning the commercial payoffs of recombinant DNA research. It is clear from the article that one of the 'fall-outs' of recombinant DNA research is the emergence of a group of scientist-entrepreneurs. Practically all the research which has led to the discovery and techniques of recombinant DNA was funded by public monies. It is obvious from the article that the recipients of these grants have jumped in with both feet in aiding to set up commercial companies to capitalise on the possible future profits resulting from the large-scale exploitation of recombinant DNA techniques. One wonders who will gain the monetary profits—only the stockholders of the companies or the scientist-entrepreneurs as well. One also wonders at the disinterestedness of the scientists, who clamoured for no restrictions on the research, giving as a reason the possible great public health benefits which might result. It seems clear now that if the benefits materialise they will accrue not only to the public, but also to the companies and the scientists. It seems to me that there is an ethical issue in all this, for while in our capitalistic society companies are supposed to make money, in our scientific society scientists are supposed to search for 'the truth', for the benefit of all and are not supposed to make money from their research, particularly if it was funded by public monies.

PHILIP SIEKEVITZ
Rockefeller University, New York, USA.

news and views

Unsuspected relatives of the ovalbumin gene

from Norman Carey

THE power of cloning techniques in the analysis of the eukaryotic genome is once again dramatically demonstrated by a paper published in this issue of *Nature* (page 125). Using the new methodology, genes apparently related to, and lying close to, the ovalbumin gene have been discovered. It seems unlikely that any other procedures would have uncovered their existence. This paper, whilst revealing yet further complexities in the organisation of the eukaryotic genome, could also lead to studies in which the control of gene expression by steroid hormones can be elucidated.

The abundant capability of the avian oviduct for protein synthesis which is under the control of steroid hormones, has lured a number of major groups to study this tissue in the hope of discovering the basis of at least one mechanism of control of protein synthesis. The early phases of this work (reviewed for example by Palmiter *Cell* **4**, 189; 1975 and O'Malley & Means *Science* **182**, 610; 1974) established that oestradiol, with or without other hormones, stimulates the growth of the tissue and the synthesis of the four major egg white proteins, ovalbumin, conalbumin, ovomucoid and lysozyme, which together can account for over 80% of the protein synthetic capacity of the cell (Palacios *et al. J. biol. Chem.* **247**, 2316; 1972). It has also been shown that the presence of the hormones promotes the accumulation of the messengers for these proteins, and if the hormones are withdrawn, the messenger content of the tissue declines to very low levels. This, and similar observations, has led to the proposition that oestrogens induce the synthesis of the proteins by a process which is the eukaryotic equivalent of the Jacob-Monod hypothesis of operon induction, a contention which has proved very difficult to support. The three effective steroids, oestradiol, progesterone and testosterone, affect the synthesis of the

major protein components differentially, in such a way that the induction process seems likely to be very complex. The experiments designed to test the proposal have been open to more than one interpretation and, in addition, other effects of steroids have been observed, such as the ability to stabilise mRNA (Palmiter & Carey *Proc. natn. Acad. Sci. U.S.A.* **71**, 2357; 1974).

This phase of primarily kinetic studies drew to a close with the advent of gene cloning. This provided a way of analysing the structure of the genes in the genome, and thus identifying the control regions which are postulated to be adjacent to the genes and are presumed to associate with polymerases and controlling proteins. The cloning of the ovalbumin gene has been dominated by the work of two groups, those of O'Malley in Texas and Chambon in Strasbourg. The first step for both was to clone the structural gene obtained from mRNA. In O'Malley's laboratory, this led first to the elucidation of the sequence of the mRNA, in collaboration with George Brownlee at the MRC Laboratory of Molecular Biology, Cambridge, and was obtained just at the time that the Fothergills in Aberdeen were putting the finishing touches to the polypeptide sequence (McReynolds *et al. Nature* **273**, 723; 1978). Chambon's group concentrated its attention on the genomic structure, and they were amongst the first to observe the presence of non-translated intervening sequences (introns) in the genes coding for polypeptides in eukaryotes (Breathnach *et al. Nature* **270**, 314; 1977; Deol *et al. Nucl. Acids Res.* **4**, 3701; 1977). The precise details of these intron/exon arrangements then rapidly became clear from the work of both Chambon's and O'Malley's groups (see for example, Garapin *et al. Cell* **14**, 629; 1978; Dugaiczyk *et al. Nature* **274**, 328; 1978; Gannon *et al. Nature* **278**, 428; 1979). The sequence which appears in the cytoplasm as messenger is interrupted in no less than seven places in the genome by DNA sequences varying from about 0.3 to 1.4 kilobase pairs. All but one of these

interruptions are found in the 5' half of the molecule which, because of the long 3' untranslated region of ovalbumin mRNA, puts them all in the coding region. The remaining intron is in the 5' untranslated region. A great deal of information has been obtained about features of the system which are shared with other genes, such as the presence of RNA precursors of messenger which are tailored to the final structure (Roop *et al. Cell* **15**, 671; 1978; Chambon *et al. Proc. 11th Miami Winter Symp.* in the press) and the nature of the sites in the RNA which are recognised by the processing enzyme (Catterall *et al. Nature* **275**, 510; 1978; Breathnach *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 4853; 1978). There are, however, other features which, while so far peculiar to the oviduct genes, must be taken into account when considering the question of eukaryotic genome organisation.

The function of introns in eukaryotic structural genes has been the subject of much speculation, and perhaps the most acceptable suggestion is that they flank regions of the gene coding for domains in the polypeptide to which portions of the total function of the protein can be ascribed. It is postulated that the genomic DNA regions coding for these domains have been brought together during evolution by a process involving the introns (Sakano *et al. Nature* **277**, 627; 1979; see *News and Views* **277**, 598; 1979). However, it is hard to see how this applies to ovalbumin. No specific functions have been ascribed to this protein. It may have important nutritional, hydrodynamic or other similar properties in relation to maintaining the developing embryo, but no function has been described or postulated which would require the cooperation of seven separate functional domains in the polypeptide. However, the function of the first intron of the ovalbumin gene must certainly be different since it lies in the 5' untranslated region of the mRNA and so has no relationship with any functional domains of the polypeptide. (Breathnach *et al. Proc. natn. Acad.*

Norman Carey is Director of Research, G. D. Searle & Co. Ltd, High Wycombe, Bucks.

Sci. U.S.A. **75**, 4853; 1978; Royal *et al.* this issue of *Nature*, page 125). Similarly, in the ovomucoid gene, a large number of exons has been found in the genomic DNA coding for a small protein which seems unlikely to have a complicated function requiring the assembly of a series of polypeptide domains (Catterall *et al.* *Nature* **278**, 323; 1979). Nevertheless, the polypeptide domain hypothesis remains an attractive one and it is possible that domain reassortment has occurred during the evolution of the oviduct proteins, but for reasons unrelated to the function of the polypeptide.

The latest piece of work in this fascinating series is published in this issue of *Nature* and is another tour de force from the powerful teams of Chambon and Kourilsky. Clearly, if there is any feature of the eukaryotic genome which resembles an operon in relation to hormonal induction, features in the DNA regions surrounding the ovalbumin gene would be expected to reveal it. Chambon and Kourilsky have therefore begun to 'walk along' the genome proceeding outwards from ovalbumin. To do so, they have used a new generation of recombinant DNA vectors called 'cosmids' which combine some properties of plasmids with the high efficiency of bacteriophage *in vitro* DNA packaging procedures, giving a very efficient selection of recombinants containing large sections of foreign DNA. Various partial digests of the chicken genome were cloned by Chambon and Kourilsky and were selected by hybridisation to a probe which covers the 5' region of the genomic ovalbumin gene, both introns and exons. Two clones are described in some detail. One, pAR1, starts within the third ovalbumin intron and extends for a total of 17 kb pairs in the 3' direction, about 12.5 kb pairs being outside the ovalbumin gene. The second, pAR2, is a 30 kb pair fragment which extends in a 5' direction from within the ovalbumin gene, and includes its first four exons. The two fragments thus cover a total of about 46 kb pairs, and overlap within the ovalbumin gene, which occupies about 7.7 kb pairs.

Certain peculiarities in the hybridisation data led the group to question whether the regions which they had cloned contained any other genes which are expressed in the oviduct. This was assessed by hybridising the total poly(A)-containing RNA from oviducts with the clones, and examining the hybrids in the electron microscope. No other genes were found on the 3' side of the ovalbumin gene, that is in the pAR1 clone, but two were found on the 5' side in pAR2. They are named X and Y, gene X being the most distal from ovalbumin, and its 5'

region seems not to be present in the clone but the RNA which hybridises to it is 2,400 nucleotides long. Gene Y RNA is about 2,000 nucleotides long, near to the size of ovalbumin mRNA, and the gene, like ovalbumin, has seven introns. The first intron of gene Y is close to the 5' end, in both X and Y all the introns are in the 5' half of the gene, and the exons of X, Y and ovalbumin are of similar size. This similarity of intron-exon pattern is not all. Parts of these genes cross hybridise with each other, and both with parts of the ovalbumin gene, which accounted for the hybridisation peculiarities which set the work on their trail in the first place. The conclusion is that there are sequence homologies in parts, but not all, of the expressed regions of all three genes.

There is also good evidence to show that the expression of all three genes is under hormonal control. The messengers are present in oviducts from birds treated with oestrogen but disappear when the hormone is withdrawn. The quantitation of this expression awaits further work, but the indications are that X and Y are only expressed to a small extent compared with ovalbumin. Although the data are not presented, the authors state that neither X nor Y codes for any of the other abundant oviduct proteins conalbumin, ovomucoid or lysozyme, but which proteins they do code for, and whether the messengers are in fact translated in the oviduct, remains unknown for the present. There is a great similarity between the organisation of the ovalbumin-related genes and that of the human β globin-related genes (see Little *et al.* *Nature* **278**, 227; 1979), where four genes (two γ genes, δ and β) of similar intron pattern are found close together on the genome. In the case of both gene sets, the 5' to 3' orientation of the genes is the same, and they are transcribed in the same direction. A further interesting feature is that the genes of each set are not expressed to the same extent under the same conditions, ovalbumin predominating over X and Y in the oestrogen-treated oviduct, and the β globin expression varying with the life cycle or disease state. We may now expect that information on the relatedness of the exons of the ovalbumin cluster will be useful in suggesting how these genes were assembled during evolution. The hybridisation data indicate that some exon regions of ovalbumin, X and Y, are closely related whereas others are more distantly related. The fine details of this relationship, perhaps including comparisons with other avian species, should indicate whether the genes have been assembled in parallel during evolution from a mixture of closely

related and unrelated regions, as the hybridisation data seem to suggest. In that case, the similarity in size of the exons may turn out to be coincidental. Alternatively, the cluster may have arisen by duplication of an ancestral gene, thus fixing the intron-exon pattern from the outset, but leaving the need to find an alternative explanation for the differing relatedness of the exons.

As for hormonal control, it has to be admitted that we are as yet no further ahead in our understanding. However, the close proximity on the genome of genes which are expressed to different extents under the influence of one hormone poses the question of whether transcription of the genes is initiated separately and with differing efficiency commensurate with the messenger abundance, or whether they are transcribed to the same extent by the same polymerase, operon fashion, with subsequent processing and stabilisation accounting for the differing levels in the cell. In other words the time may now be approaching when workers in the field can return to the study of the kinetics of the induction process with some real hope of testing the hypotheses. □

Histocompatibility antigens and cytoskeletal elements

from Franklin H. Portugal

MAJOR histocompatibility antigens designated as H-2 in mice and HLA in man are cell-surface glycoproteins whose movement in the plasma membrane appears to be associated with the movement of intracellular cytoskeletal elements. These antigens are involved in such immune reactions as allograft rejection and recognition of viral and tumour antigens by syngeneic cytotoxic thymus-derived (T) lymphocytes. Initially, it was difficult to isolate even a few micrograms of pure HLA antigen from human splenic lymphocytes. Observations that the surface of cultured human lymphoblastoid cells contains a greatly enhanced antigen concentration simplified this problem and resulted in the isolation of milligram amounts of material. In 1975, Peterson *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **72**, 1612; 1975) proposed that the large chain for H-2 and HLA antigens, already known to carry the alloantigenic determinants, has both

Franklin H. Portugal is a science writer at present working on a history of the Carnegie Institution of Washington.

chemical and physicochemical properties similar to those of immunoglobulin. These antigens also contain a smaller chain which consists of β_2 -microglobulin, is invariant, and shows homology to various IgG heavy and light chain homology regions. Because major histocompatibility antigens are found on the surface of virtually all normal, nucleated cells in man and mice, probably all mammalian cells have an immunoglobulin-like molecule on their surface. By solubilising H-2 antigen with either a proteolytic enzyme (papain) or detergent, Henning *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **73**, 118; 1976) showed that the amino-terminal end of the heavy chain extends away from the cell surface whereas the carboxyl-terminal end is associated with the cell membrane. Further characterisation indicated that H-2 antigen consists of a light-chain β_2 -microglobulin (12,000 daltons) and a heavy polypeptide chain (46,000 daltons) that carries both distinctive antigenic determinants as well as a carbohydrate prosthetic group (Edelman *Science* **192**, 218; 1976). Similar studies with HLA-A antigen showed that it contained a light chain of β_2 -microglobulin (12,000 daltons) and a heavy polypeptide chain (44,000 daltons) containing both antigenic specificity and carbohydrate (Springer & Strominger *Proc. natn. Acad. Sci. U.S.A.* **73**, 2481; 1976). Springer and Strominger proposed that a hydrophobic region of the heavy chain is inserted into the cell membrane so that the amino-terminal portion of the chain extends outside the cell and the carboxyl-terminal portion extends into the cell interior. By using the lactoperoxidase-catalysed iodination of the inner surface of lymphocyte plasma membrane, Walsh and Crumpton labelled HLA and showed therefore that this antigen is a transmembrane protein (*Nature* **269**, 307; 1977).

How the structure of these antigens and their insertion in the cell membrane may affect immune reactions can be studied with liposomes. Littman *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **76**, 902; 1979) have prepared unilamellar phosphatidylcholine liposomes containing H-2 antigen. In these liposomes, H-2 antigen was found to be integrally inserted in the membrane, oriented with the amino-terminal end of the chain facing away from the cell surface, and antigenically active. Such antigen-containing particles can also be used in assays with T lymphocytes.

Antigen movement over the cell surface is detectable by patching and capping phenomena. Bourguignon and Singer used a double fluorescence technique for staining both surface-bound H-2 antigen on mouse splenic T lymphocytes and for staining intracellular mechanochemical proteins such as actin or myosin (*Proc. natn. Acad. Sci.*

U.S.A. **74**, 5031; 1977). When the surface-bound receptors collected in patches in response to an external multivalent ligand, the membrane-associated actin or myosin within the cell also accumulated in patches directly under those of the receptor. This suggests that H-2 glycoprotein becomes linked to the intracellular contractile protein, possibly through an integral class of proteins that are located in the plasma membrane of eukaryotic cells. Koch and Smith confirmed that there is an association between H-2 antigen and actin (*Nature* **273**, 274; 1978), although the question of whether the association is a direct or indirect one could not be resolved.

Pober *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **75**, 6002; 1978) have recently indicated how such an association between the intracellular domain of a transmembranous protein and the cytoskeletal protein may be regulated. Incubation of either transformed lymphoblastoid cells or peripheral blood lymphocytes with $^{32}\text{P}_i$ resulted in the phosphorylation of HLA-A and HLA-B antigens. The site of phosphorylation

is a serine residue(s) located in the hydrophilic, carboxyl-terminal region of the heavy chain. Incubation of isolated membranes from transformed lymphoblastoid cells with $[\gamma^{32}\text{P}]\text{ATP}$ also produced phosphorylation of HLA antigens. Because interaction with actin is known to occur with many proteins that are phosphorylated, the phosphorylation of HLA antigens may be the basis for regulating association with cytoskeletal elements.

Many questions remain to be answered in future experiments. One concerns what factors regulate the phosphorylation of these membrane proteins. The question of which serine residue(s) of HLA-A and HLA-B antigens are phosphorylated and the number of phosphate groups incorporated into each heavy chain requires further resolution. Presumably, HLA antigen binds to actin as does H-2 antigen but this also remains to be demonstrated. It is still unclear whether microfilaments or microtubules or both are the cytoskeletal elements involved in the association with transmembranous proteins. □

Laser spectroscopy pins down Rydberg constant

from John E. M. Goldsmith

THE availability of tunable laser radiation and associated nonlinear spectroscopic methods has revolutionised many areas of spectroscopy (for a review, see *High-Resolution Laser Spectroscopy* Ed. Shimoda.) Springer-Verlag, Berlin, 1976). One notable example is the elimination of Doppler broadening in the observation of the Balmer-alpha line ($n=2 \rightarrow n=3$, $\lambda=656\text{ nm}$) of atomic hydrogen with the technique of saturated absorption spectroscopy (Hansch, Shahin & Schawlow *Nature* **235**, 63; 1972). The Doppler effect causes atoms moving with different speeds to appear to have different resonance frequencies, just as the apparent pitch of a train whistle changes as the train moves past the observer. Since the atoms in a gas have a wide range of velocities, and all of the atoms are observed using classical (linear) spectroscopic techniques, the net effect of the Doppler shift is to cause a broadening of the spectral line. With saturated absorption spectroscopy, only atoms that are not moving (or at least have no component of velocity along the laser beams) are observed, thus eliminating Doppler broadening. The increase in resolution obtained in this

manner makes it possible to observe the individual fine structure components of the Balmer-alpha transition.

The ability to resolve this fine structure has great importance for measurements of the Rydberg constant. The Rydberg, which is defined as an explicit combination of other fundamental constants $R = \mu_0^2 m_e e^4 / 8h^3$, is a cornerstone in the evaluation of the fundamental physical constants. The physical meaning of the Rydberg is most clearly seen in the simple Bohr theory, in which the energy levels of the hydrogen atom are given by the expression $E_n = -hcR/n^2$, $n=1,2,3, \dots$; the principal quantum number n labels the energy levels of the atom. Thus the value of the Rydberg constant corresponds to the energy required to ionise the ground state hydrogen atom.

Until 1974, the Rydberg constant was determined by measuring the wavelength of emission lines from cooled hydrogen, deuterium, tritium, or ionised helium discharges (see *Series, Contemp. Phys.* **14**, 49; 1974, for a review of these measurements). The fine structure components of the lines were not resolved because of Doppler broadening, so it was necessary to calculate the spectral lineshape produced by the blend of components. This caused an uncertainty in pin-

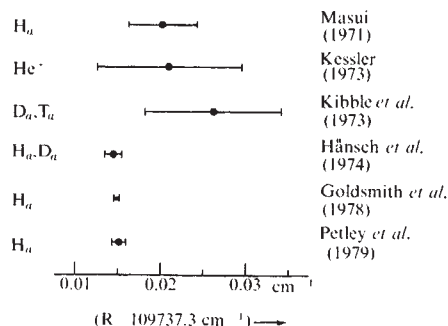
John E. M. Goldsmith is in the Department of Physics, Stanford University.

pointing the 'centre' of the emission line, and thus an uncertainty in the value of the Rydberg constant.

Resolution of the fine structure of the Balmer-alpha line with Doppler-free laser spectroscopy has made possible nearly two orders of magnitude improvement in the precision of the Rydberg constant, making it one of the (if not the) most precisely measured of all fundamental physical constants. The first Doppler-free measurement used saturated absorption spectroscopy with a pulsed dye laser to observe the hydrogen and deuterium $2P_{3/2}-3D_{5/2}$ transitions. A d.c. discharge tube similar to Wood's original design was used to dissociate the hydrogen molecules and excite some of them to the $n=2$ level. The Rydberg constant was determined from a measurement of the wavelength of this transition, using an iodine stabilised helium-neon laser ($\lambda=633$ nm) as a wavelength standard. The value so obtained, $R=109737.3143(10)$ cm^{-1} (Hänsch *et al.* *Phys. Rev. Lett.* **32**, 1336; 1974), was an order of magnitude more precise than previous Doppler-broadened measurements.

Further improvement in the Rydberg constant has been aided by the development of narrowband continuous wave (cw) dye lasers, and suitable dyes for cw operation at the Balmer-alpha wavelength. Using a variation of saturated absorption spectroscopy (known as polarisation spectroscopy) with a cw dye laser, and replacing the Wood's tube with a mild helium-hydrogen discharge, it became practical to study the hydrogen $2S_{1/2}-3P_{1/2}$ transition, which is weaker than the $2P_{3/2}-3D_{5/2}$ transition, but also more than three times narrower. The wavelength measurement of this transition was facilitated by first determining the wavelength of an iodine reference line which by good fortune has a wavelength almost identical to that of the $2S_{1/2}-3P_{1/2}$ transition. A measurement of this small wavelength difference then yielded the value $R=109737.31476(32)$ cm^{-1} (Goldsmith, Weber & Hänsch *Phys. Rev. Lett.* **41**, 1525; 1978), with a precision three times higher than obtained previously. A more precise measurement of the wavelength of this iodine reference line could immediately yield at least another twofold improvement in the precision of the Rydberg constant.

Another recent Rydberg measurement returned to the use of saturated absorption spectroscopy in a Wood's discharge tube, but with a narrowband cw dye laser. The wavelength of three hydrogen Balmer-alpha fine structure components were measured ($2P_{3/2}-3D_{3/2}$, $2S_{1/2}-3P_{3/2}$ and $2P_{1/2}-3D_{3/2}$), with the iodine-stabilised helium-neon laser again used as a wavelength



standard. The weighted mean of these values yielded the result $R=109737.31513(85)$ cm^{-1} (Petley and Morris, this issue of *Nature* page 141).

The values and precision of Rydberg measurements from the past 10 years are summarised in the figure. The Doppler-free measurements clearly are much more precise than the earlier Doppler broadened measurements. It is more significant, however, that these three independent measurements agree very well with each other, but are only marginally consistent with the earlier determinations. The possibility of such a systematic error indicates the need for independent measurements of values such as the Rydberg constant. The excellent agreement of the three Doppler free determinations, performed in two laboratories on several Balmer-alpha fine structure components using varied techniques, is a strong indication that such a systematic error has been avoided.

Measurement of the Rydberg constant at the current level of precision is already hindered by uncertainties inherent in the ^{86}Kr primary wavelength standard, which limit its useful precision to a few parts in 10^9 . This is an artificial difficulty, however, and can be avoided by using lasers stabilised to molecular transitions, such as the iodine-stabilised helium-neon laser, as interim wavelength standards. Direct frequency measurements have recently been extended into the infrared region of the spectrum, and frequency measurements in the visible may not be too far off. For measurements with a precision greater than 1 part in 10^9 it will probably prove necessary to work with an atomic hydrogen beam, so small energy level shifts caused by conditions in the discharge can be eliminated. The ultimate Rydberg measurement could come from a frequency measurement of a narrowband cw ultraviolet laser source ($\lambda=243$ nm, $\nu=10^{15}$ Hz) used to excite the $1S-2S$ transition in hydrogen using Doppler-free two-photon spectroscopy (Hänsch *et al.* *Phys. Rev. Lett.* **34**, 307; 1975). The natural linewidth of this transition is about 1 Hz, so one

can conceive of a measurement of its frequency with a precision better than 1 part in 10^{15} . Using modern atomic theory, it is necessary to apply several small corrections to the hydrogen energy levels calculated with Bohr's theory. With such a frequency measurement, the precision of the Rydberg constant would only be limited by the precision with which these small corrections can be calculated. □

Ionic landmarks along the mitogenic route

from K. S. Koch and H. L. Leffert

Two ionic fluxes, akin to those initiating and transducing nerve impulses, may be essential regulators of animal cell proliferation. This idea emerged at a recent conference* where many findings pointed to increased Na^+ influx being an early, perhaps the initial, mitogenic signal, and to subsequent Ca^{2+} ion movement coupled to cyclic AMP formation being necessary for DNA replication.

Experiments using amiloride (a specific Na^+ influx inhibitor) and $^{22}\text{Na}^+$ showed that 'burst-like' influxes are needed to initiate DNA synthesis in sea urchin eggs after fertilisation (D. Epel, Stanford University), in rat hepatocytes *in vitro* after exposure to peptide growth factors including EGF (S. Cohen's epidermal growth factor) (K. S. Koch & H. Leffert, The Salk Institute), and in cultured mouse 'fibroblasts' after exposure to serum factors including vasopressin (E. Rozengurt, Imperial Cancer Research Fund, London). Similar fluxes appeared operative during phytohaemagglutinin-induced human lymphopoiesis (G. Kaplan, University of Ottawa) and veratridine-stimulated mitogenesis of cultured chick central nervous system neurones (C. D. Cone, Hampton Virginia VA Hospital).

Possible mechanisms by which growth factors stimulate Na^+ influx were considered. G. Eisenman (University of California, Los Angeles) discussed current views of cation-selective channel functioning based on gramicidin models and proposed that certain mitogens directly form the Na^+ channel—an idea awaiting experimental test. Indirect mitogenic actions on channel 'gating' mechanisms result-

*'Growth Regulation by Ionic Fluxes' was held on 13-17 March 1979 at the Kroc Foundation, Santa Ynez, California, and the proceedings will be published as a number of the *Annals of the New York Academy of Sciences*.

ing from altered surface potentials (C. Bergman, University of Paris); upon lateral diffusion of membrane proteins resulting from altered membrane potentials (M. Edidin, Johns Hopkins University) measured with lipophilic cations like triphenylphosphonium⁺ (R. Kaback, Hoffman-LaRoche); and on cell-cell interactions resulting from altered electrical coupling (W. Loewenstein, University of Miami) were also discussed.

How might increased Na⁺ influx stimulate proliferation? One attractive mechanism, generally consistent with the experimental results described, may be activation by intracellular Na⁺ of the membrane Na⁺/K⁺ ATPase 'pump'. Such changes are expected to alter Na⁺ gradients across the surface membrane. The coupling of solute transport to these gradients through a membrane 'carrier'—presumably to stimulate biosynthetic pathways—was reviewed in detail by R. Crane (Rutgers University). He showed evidence for amiloride-sensitive Na⁺ gradient-dependent glucose uptake into rabbit intestinal vesicles as well as reconstitution of this system with a purified protein, obtained from brush-border membrane fractions, inserted into liposomes. The effects of growth factors in Crane's vesicle systems are unknown. Another possible consequence of increased Na⁺ influx studied in fertilised eggs—a Na⁺/proton exchange elevating intracellular pH—generated a lively debate between its proponent (Epel) and its antagonist (L. F. Jaffe, Purdue University). Epel speculated that small pH changes cause widespread alterations involving critical enzyme cascades and macromolecular polymerisations. Jaffe then described egg experiments where injected buffers that blocked alkalisation failed to block growth. F. M. Harold (National Jewish Hospital, Denver) added that disruption of proton gradients with carbonylcyanide m-chlorophenylhydrazone in bacterial systems failed to impair growth seriously. But the issue was not resolved, and the role of Mitchell's proton gradients as regulators of animal cell proliferation remains to be explored. Further consequences of increased Na⁺ influxes were suggested by J. Piatigorsky (National Institutes of Health) and G. Koch (University of Hamburg) whose results with ocular lens and viral protein regulatory systems, respectively, raised the possibility of Na⁺ and K⁺ selective translational and/or post-translational controls. These inquiries may soon yield definitive conclusions, especially

if a factor isolated by G. Koch, a dialysable methionine-containing translational inhibitor released from ribosomes by high Na⁺, is purified and if Piatigorsky's applications of recombinant DNA technology are successful.

A fourth likely consequence of increased Na⁺ influx was elegantly demonstrated by Jaffe who showed by the use of aequorin, a Ca²⁺-sensitive photoluminescent protein, the flow of a Ca²⁺ 'wave' across the cytoplasm of a fertilised Medaka egg. If this 'explosive' wave was blocked by intracellular injections of Ca²⁺-chelators, DNA synthesis did not begin. Jaffe theorised that Ca²⁺ 'currents', detected extracellularly by a 'vibrating' probe developed earlier in his laboratory, cause self-electrophoresis of membrane proteins for which evidence was presented (K. Robinson, National Jewish Hospital). Whether such Ca²⁺ currents flow across somatic cells stimulated to divide is not yet known. However, a clear late G_i-requirement for surface Ca²⁺ movement was demonstrated—for a broad spectrum of normal cells *in vivo* and *in vitro*—without which deoxyribonucleotide, and consequently DNA, synthesis cannot occur (A. Boynton, NRC, Ottawa). These events are mediated by cyclic AMP and it was postulated that cellular ribonucleotide reductase function requires a phosphoprotein regulator whose formation depends on cyclic AMP-activated protein kinase.

Properties of nervous system and limb regeneration were reviewed by R. Moore (University of California, San Diego) and by Jaffe, respectively. An attempt was made to consider these processes with regard to the ionic events described above. Two provocative links were provided, first from Jaffe's evidence that chronically applied cathodal Na⁺ currents stimulate partial limb regeneration in frogs and second, from F. Westall's (The Salk Institute) observations that proteolysis converts myelin basic protein into a mitogenic peptide whose sequence (Thr-Pro-Pro-Ser-Gln-Gly-Lys) is contained within the brain fibroblast growth factor (FGF) isolated by D. Gospodarowicz (University of California, San Francisco). Such findings immediately provide one explanation for nerve requirements in limb regeneration. Moreover, although Gospodarowicz showed that myoblasts, chondrocytes and fibroblasts proliferate in response to FGF, he argued that blastemal cells, from which regenerated tissue arises, may be this mitogen's natural target. Thus, Jaffe's Na⁺ 'currents' in regenerating stumps might actually be driven by FGF-like mitogens released from crushed myelinated nerves.

The existence of two regulatory ionic

'landmarks' has interesting implications for understanding growth regulation in general. In animal cell systems, it helps to explain how sets of growth factors may act synergistically for it predicts that certain substances activate Na⁺ influx while others act predominantly upon Ca²⁺-cyclic AMP coupling. Reason to believe that two-signal models are correct was presented for hepatocytes (K. S. Koch & Leffert; Boynton) and has been reported elsewhere for BALB/c 3T3 cells (see for example Stiles *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 1279; 1979). Furthermore, interference with the Na⁺ flux system perturbs transitions from resting to growing states differently from interference with the Ca²⁺-cyclic AMP couple. The former seems to determine both the lag-duration (or 'onset' time) and the total numbers of responding cells (K. S. Koch & Leffert; Rozengurt) whereas the latter mechanism seems to regulate the rate of entry into S-phase (Boynton). These findings must be considered by any theory of growth regulation; single-event models are insufficient to account for such observations. Lastly, it seems that control of both events is defective in tumour cells (Cone; Rozengurt; Boynton) but the real significance of these observations remains to be determined. □

Studying surfaces

from M. A. Chesters

RECENT progress in understanding the interaction of molecules with solid surfaces was a major theme of the ECOS2 Conference at Cambridge*. Although the subjects covered included adhesion, Auger microanalysis and the chemistry of polymer surfaces, about 70% of the papers were concerned with chemisorption on metal surfaces. D. Tabor (Cavendish, Cambridge) in his opening address traced the history of surface science and urged that increased attention be paid to fundamental studies of liquid/solid and solid/solid interfacial phenomena. As it was the conference provided an excellent review of progress in fundamental studies of chemisorption, surface elemental analysis and surface structural analysis.

A major motivation for studies of chemisorption is the need to understand heterogeneous catalysis on a molecular level and two approaches to the problem of linking research on idealised systems to the behaviour of

K. S. Koch and H. L. Leffert are in the Cell Biology Laboratory, The Salk Institute, San Diego, California.

*The Second European Conference on Surface Science was held at the Cavendish Laboratory and Churchill College, Cambridge, England, on 26–29 March 1979.

real catalysts were described by G. A. Somorjai (University of California, Berkeley) and R. J. Madix (Stanford University). Somorjai has pioneered work on high index (stepped) metal single crystal surfaces in an attempt to identify the types of active site involved in various catalytic processes. He described experiments carried out in an ultra-high vacuum system combined with high pressure cell, enabling metal single crystal catalysts to be characterised by low energy electron diffraction and Auger spectroscopy before and after catalysing reactions at pressures up to 200 atmospheres. Apart from this direct approach, molecular beam studies of reactive scattering from stepped surfaces revealed the importance of step and kink surface sites in breaking carbon-hydrogen and carbon-carbon bonds. Madix argued that surface intermediates important in reactions catalysed at high pressure and high temperature, may be isolated in chemisorption studies at low pressure and low temperature. He described the use of thermal desorption spectroscopy coupled with photoelectron spectroscopy to identify a surface formate intermediate in the decomposition of formic acid on Ag(110) and a surface methoxy intermediate in the decomposition of methanol on Fe(100).

In the measurement of surface reaction kinetics, molecular beam scattering remains the major technique and was reviewed by G. Ertl (München University). G. W. Rubloff (IBM, New York) described a rather more direct approach to the study of rapid surface reactions in which surface concentrations were monitored by ultraviolet photoelectron spectroscopy (UPS). The advantage of such a technique is that surface reactions which do not involve desorption of gas may be studied. Both B. I. Lunqvist (Chalmers University of Technology, Sweden) and T. B. Grimley (University of Liverpool) outlined theoretical formalisms for describing the dynamics of the interaction of atoms and molecules with solid surfaces which should pave the way for the calculation of sticking coefficients and the construction of chemisorption potential energy curves. The increased experimental and theoretical interest in studying the kinetics of reactions on well defined surfaces was particularly apparent at the conference. At the same time, there was an encouraging, continued trend away from purely 'technique-oriented' discussions, although this cannot be completely avoided in an area where new techniques abound.

The formidable problem of the elucidation of the geometrical and electronic

structure of metal-adsorbate complexes was discussed largely in the context of photoelectron spectroscopy. G. J. Lapeyre (Montana State University) illustrated how the operation of the symmetry selection rule in UPS is revealed in angular resolved experiments using polarised radiation (PARUPS), thus allowing the symmetry of the ionised states to be determined. In a discussion of the requirements of a full theoretical analysis of photoemission from adsorbates, J. B. Pendry (SRC Daresbury Laboratory) remarked that if the theoretical analysis were straightforward we would learn little from the experiment! P. H. Citrin (Bell Laboratories) later reviewed the SEXAFS experiment in which the fine structure above an X-ray absorption edge, which results from diffraction of the photoelectron, may be analysed on a straightforward kinematical basis to yield accurate data on the relative positions of atoms surrounding the photoelectron emitter. Surface Extended X-ray Absorption Fine Structure is detected through modulations in the Auger yield from the absorbing atom to achieve surface sensitivity, and Citrin illustrated the power of the technique by reporting results for iodine adsorbed on silver and copper single crystal surfaces. The chief disadvantage of SEXAFS is that a synchrotron radiation source is essential, so we shall be hearing much more of the technique as storage rings dedicated as photon sources become operational.

In the past few years there has been a resurgence of electron energy loss spectroscopy (EELS) as a means of studying vibrational modes of adsorbed species. This is in large measure attributable to H. Ibach (KFA, Jülich) and he gave an account of its application to the study of hydrocarbon adsorption on nickel surfaces. The technique involves the measurement of small energy losses (40–500 meV) of electrons impinging on a solid surface which arise through excitation of vibrational modes of the surface and of adsorbed molecules. The sensitivity is about two orders of magnitude higher than for reflection absorption infrared spectroscopy. Ibach was able to show some striking differences between the reactivity of a flat Ni(111) surface and a stepped Ni[5(111)×(110)] surface towards acetylene and cyclohexane. R. F. Willis (ESTEC, The Netherlands) has carried out a detailed investigation of the W(100)/hydrogen system in which he has pioneered the use of EELS in both specular scattering and diffuse scattering modes for which different selection rules apply. He described a detailed model for a reconstructed surface in which tungsten-hydrogen-tungsten bridges are tilted

out of the surface plane. The apex angle of the bridge and the tilt angle were determined from the frequencies and angular dependences of the three hydrogenic vibrational modes. Clearly EELS can supply some of the basic structural information on adsorbates which has proved so elusive and will be a vital complement to the other electron spectroscopic techniques in the future.

There is a continuing debate on the relevance of these idealised investigations to 'real' problems in surface science (see Tompkins, *Chemistry in Britain* 15, 194; 1979). For myself the conference was preaching to the converted. I leave others to make their own judgement. □

Birds as agricultural pests

from John Krebs

THE male bullfinch not only looks beautiful; it can be trained to whistle German folk tunes, and given the chance it will strip half the fruit buds off a pear tree in 90 minutes. It was this last attribute which brought the bullfinch to the attention of a recent meeting on birds as agricultural pests.* J. M. Flegg (East Malling Research Station) reviewed the damage caused by bullfinches, redpolls and other birds to fruit crops, and pointed to some common sense factors influencing the level of destruction. Orchards planted next to natural woodland are more likely to suffer bullfinch damage as the woodland provides a good home base for the birds. Alder trees planted as windbreaks draw in redpolls to feed on the cones, and when the cone supply is exhausted the birds switch to devouring fruit buds; and the habit of removing all the herbaceous weeds from orchards may increase bullfinch damage, since the bullfinch prefers weed seeds and only switches to fruit buds when seeds are not available.

Farmers have tried to alter the food preferences of bird pests for hundreds of years. Although older methods such as rubbing garlic or pigs' dung on fruit trees are now outmoded, a modern equivalent is to try and condition birds to dislike crop foods by treating the crop with a chemical which makes the birds feel sick. This kind of aversive conditioning is well known to induce

*The Symposium on 'Understanding Agricultural Bird Problems' was organised by the British Crop Protection Council and Agricultural Science Service at Royal Holloway College, 4–5 April 1979. The proceedings are to be published (by the British Crop Protection Council) as a book edited by E. N. Wright, I. Inglis & C. J. Feare.

M. A. Chesters is in the School of Chemical Sciences, University of East Anglia, Norwich.

'bait shyness' in rats and J. Rogers (US Fish and Wildlife Service) reported that it also works with birds. In laboratory trials, red-winged black-birds would avoid eating a foodstuff for more than 16 weeks after first encountering it mixed with the insecticide methiocarb, and this conditioned aversion was effective even when the only alternative food was highly unpalatable. Large scale field trials showed that methiocarb can reduce bird damage to vineyards by something like fivefold.

I. Inglis (Pest Infestation Control Laboratory, Worpleston) described the efficacy of different kinds of visual scaring devices. The traditional scarecrow, and popular modern equivalents consisting of bright coloured miniature windmills are totally ineffective, and Inglis suggested that a knowledge of the behaviour of pest species is useful in designing more effective bird-scarers. For example, when a wood-pigeon is frightened and takes to the air, it reveals conspicuous white wing patches which might be seen by other pigeons as a danger signal. This led Inglis to try out the deterrent effect of pairs of pigeon wings lying on the ground, and in one study birds were largely deterred from landing in a clover field for up to 70 days by the sight of pigeon wings on the ground. Similarly, Brent geese were discouraged from landing in a field by putting out model geese in an 'alert anxiety' posture. One of the most effective deterrents of all is the sight of a man flapping his arms up and down like a giant eagle, and there is now available on the market a 7-foot high Incredible Hulk with mechanical flapping arms. All these scaring devices suffer from the problem of habituation: birds eventually learn to stop flying away from them unless occasionally reminded with a gun that there is a real danger. It may turn out in the end, as suggested in one of the discussions, that the best solution to bird scaring is the one recommended in 1668 by Gervais Markham in *Farewell to Husbandry*: "The only best and safest means to prevent this evil is to have eversome young boys with bows and arrows to follow the seedman and harrows, making a great noise and shooting his arrows where he shall see these devourers aflight, not ceasing but chasing them from the land".

The meeting also touched on a more fundamental ecological issue raised by all techniques of scaring, deterring, or luring away bird pests: what happens to birds when they are frightened off a particular field? If they go next door, the problem is merely shifted on to someone else. One possibility, discussed by M. Owen (Wildfowl Trust, Slimbridge) is to provide refuges for birds,

The leaning tower

from Ian Smalley

BECAUSE the Tower of Pisa was built very slowly, with many interruptions, it still stands; had it been completed during the 12th century it would probably have collapsed (Mitchell, Vivatrat & Lambe *J. Geotech. Eng. Div. Amer. Soc. Civ. Eng.* **103**, 227; 1977). Work on the foundations of the tower began in August, 1173 and the first story was completed in 1174. When the tower reached a height of three and a half storeys and a load of 9,840 metric tons in 1178 the work stopped. The work stoppage has been attributed to politics, to the heavy work load of the constructors, and to construction difficulties. Construction was not resumed until almost a century later and then the tower was completed up to the eight storey level, and to a total load of 13,728 t, during the period 1272-1278. Work then stopped again and did not resume until 1360 when the final storey was added and the whole tower completed in 1370. By the time of the final stage of construction the lean of the tower was significant and the centre line of the topmost part was changed on account of this.

Mitchell *et al.* concluded that the bearing capacity of the soils underlying the tower was never exceeded. The total settlement of the tower is due to the sum of four components: an immediate compression of the sands in a 7 m thick zone underlying the base of the tower, immediate compression of a 30 m thick clay layer underlying the sands, consolidation of the clay layer and secondary compression of the clay layer. They were unable to show, using their classical soil mechanics methods, why the tower actually leans although they suggested that this might be due to the higher compressibility of the foundation sand on the south side of the tower. Veder (*Bauingenieur* **50**, 204; 1975) had suggested earlier that the stone blocks used in the construction were stored on the south side of the tower causing an asymmetric loading.

Although the tower leans to the south now, it seems to have pointed

in various directions during its construction. From 1173 to 1178 two-thirds (in terms of weight) of the tower was constructed and it tilted towards the north-east. By the time construction resumed in 1272 this inclination had doubled. Inclination increased for a further 6 years at which time 86% of the final weight was in place. The tower then leant towards the north-west, and in 1370 when the tower was completed its inclination was towards the south, as it is today. The tilt was 1/31 in 1370 but today it is around 1/10 and, after a deceleration of movement for five and a half centuries, movements are today tending to accelerate. Much of the recent movement can be related to the removal of water from nearby wells and the lowering of the piezometric head in the deep sandy layer (Mascardi *J. Geotech. Eng. Div. Amer. Soc. Civ. Eng.* **104**, 299; 1978).

Cambefort (*J. Geotech. Eng. Div. Amer. Soc. Civ. Eng.* **104**, 156; 1978) has calculated that the tower will be stable until the year 2780, as long as the pumping from the deep aquifer does not increase beyond the 1966 rate. He has also proposed an ingenious lever mechanism to halt the tilt and fix the tower in its present position; a lever arm extending 12.3 m from the tower centre and a load of 600 t would apparently provide the necessary counterforce. Alternative solutions to the engineering problem are however still being sought; and there is still no consensus on what actually caused the tilt. Perhaps the most ingenious theory is still that advanced by Kerisel (*Geotechnique* **25**, 433; 1975) who invoked the Coriolis force due to the rotation of the Earth. Is it possible, wondered Kerisel, that a structure with a very low safety factor might be influenced by sustained small forces acting over several centuries?

Ian Smalley is with the Soil Engineering Section, DSIR Soil Bureau, Lower Hutt, New Zealand.

and an alternative is the one mentioned earlier in the context of bullfinches, to provide a different food supply. These courses of action may, however, exacerbate the problem. If a

John Krebs is a Lecturer in Zoology at the Edward Grey Institute of Field Ornithology University of Oxford.

pest species survives better as a result of the provision of refuges or alternative food supplies, it may return in great numbers in future years to attack agricultural crops. This kind of problem cannot be resolved without more knowledge of the factors determining survival and movements of bird pests. □

One or more Eemian interglacials?

IN a recent paper (*Nature* 277, 189; 1979) we described the Fjøsanger (interglacial) Stage in western Norway, and correlated it with deep sea oxygen isotope stage 5e and the Eemian of continental Europe. Bowen (*News & Views* 277, 171; 1979) accepted the former correlation but criticised the latter, because he agrees with Kukla's proposal (*Earth Sci. Rev.* 13, 307, 1977) that the Eemian represents three different interglacials.

The vegetational development in different interglacials may have been so similar that a separation on palynological evidence is impossible. However, this is not necessarily so, and the crucial point in the present discussion is whether other interglacials repeated the distinct pollen sequence so far being considered as the criterion for the Eemian.

We fully agree with Bowen's statement that the deep sea oxygen isotope stratigraphy proves that the classical subdivisions of the Quaternary of Europe are incomplete. That is our challenge! However, an incomplete continental stratigraphy does not imply that all previously defined units are ambiguous.

Bowen claimed that Weigank (*Geologie* 21, Beiheft 77, 1; 1972) has described two Eemian interglacials in superposition in northern Germany. On the contrary, however, he stated that these two interglacials can be identified on the basis of their pollen stratigraphy, the younger being the Eemian, and the older the Rügenian (=Kap Arkona in Frenzel's classification).

Most important Bowen stated that Kukla (*op. cit.*) has demonstrated that the classical Eemian sites in Europe were deposited during three different interglacials. No doubt, Kukla made an extremely interesting analysis of the Quaternary stratigraphy of Europe, but we cannot accept his evidence for more than one Eemian. Chiefly on the basis of geomorphological observations he postulated that the Saalian s.l. is composed of three distinct glacials, and he used Eemian sites for the presumptive interglacials so created. Most West German geologists (Duphorn *et al.* *Eiszeitalter u. Gegenw.* 23/24, 222; 1973) reject the existence of an interglacial within the Saalian s.l., but

there are other more probable candidates (Rügen, Dömnitz/Wacken) to fill the milder intervals (see Weigank *op. cit.*; Erd *Palaeogeogr., -clim., -ecol.* 8, 129; 1970; Frenzel *Eiszeitalter alter u. Gegenw.* 23/24, 321; 1973; Menke & Behre *Eiszeitalter u. Gegenw.* 23/24, 251; 1973; Cepek Ber. *deutsch. Ges. geol. Wiss., A. Geol. Paläont.* 12, 375; 1967) than the Eemian, even if Kukla's interpretation is correct.

The main evidence in support of our thesis that the typical Eemian pollen sequence represents only one interglacial can be summarised as follows. Many known Eemian sites are located inside the Weichselian ice border; all of them are situated below till or disturbed by ice. More than 100 Eemian sites are known outside the Weichselian ice border, and, as far as we know, not a single sequence is reported to be below till, or below another interglacial. However, many 'not Eemian interglacials' are known below till outside the Weichselian ice border.

We conclude that our present knowledge strongly suggests that the Eemian pollen sequence of Europe represents only one interglacial, which is the Last Interglacial.

JAN MANGERUD

EIVIND SØNSTEGAARD

HANS-PETTER SEJRUP

Department of Geology,

Allégt. 41,

5014 Bergen-Universitetet,
Norway.

D. Q. BOWEN REPLIES: The points raised by Mangerud *et al.* are only valid in so far as the chronostratigraphy advocated by Kukla and others is incapable of direct validation at present. But the fact remains that the system they advocate and defend in the case of 'Eemian' is based on assemblage floras, a biofacies basis inherently unreliable as a result of its diachronous nature. Indeed by advocating that Woillard's (*Quaternary Res.* 9, 1; 1978) St Germain Interglacials, I and II, are time-equivalent to interstadials in the Netherlands and Denmark, also urged by Wijmstra (in *Climatic Change* (ed. Gribbin) Cambridge, 1978), they tacitly acknowledge the fragility of this bio-

facies approach.

Discrimination of specific interglacials on the basis of either being covered by till or not in relation to postulated glacial limits is no improvement on 'count from the top' methods criticised by Kukla. In this way the current conventional wisdom, or model chronostratigraphic schema, is maintained. Data are not only interpreted on such model terms but are frequently acquired and systematised according to its pigeon-holed ready-made classification—the 'reinforcement syndrome' of the late Norman Watkins (*Comments on Earth Sci. Geophys.* 2, 36; 1971). In the case of Weigank's work the significant point relates to his demonstration that the foraminiferal faunas of the Eemian and Rügenian Interglacials show only minor differences.

The fact is that, according to present interpretation (Shackleton & Opdyke *Quaternary Res.* 3, 39; 1973) and defining the base of the Middle Pleistocene at the Brunhes-Matuyama boundary (Butzer *Quaternary Res.* 4, 136; 1974), some eight interglacials occurred from that time to the present. It is pertinent to note that one of the UK working parties on International Geological Correlation Programme Project 24 is examining evidence to determine how many interglacials there have been since the Hoxnian. Conventionally there is only one—the Ipswichian (Eemian)! In the search for greater complexity in continental records it is inevitable that some degree of data manipulation will occur initially because many of the basic facts regarding Pleistocene geology are still inadequately known. For too long has it been assumed that the Pleistocene geology of most formerly glaciated areas is known—perhaps it is, at least according to the terms of existing models. But in terms of modern lithostratigraphic standards (Hedberg Int. Stratigraphic Guide, New York, 1976) this work, with some notable exceptions (Willman & Frye *Illinois Geol. Surv. Bull.* 94, 1970), has hardly commenced. In the meantime, other than by rare geochronometric dating, progress will inevitably include some manipulation born out of an expediency dealing with inadequate data.

A hundred years ago

WE need not insist on the extreme importance and interest of the exhibition which was opened last night at the Albert Hall, and for which extensive preparations have been making for

some time. The public mind both in this country and abroad has been recently much agitated on the question of electric lighting, and, as might be expected, people are much confused among the many systems which have been brought forward, and even those

who know something on the subject must find it difficult to make up their minds. Hence the importance of bringing together the various systems of electric lighting in such a way as to make comparison possible.

From *Nature* 20, 8 May, 39; 1879.

review article

Facts and hypotheses of molecular chemical tunnelling

Vitalii I. Goldanskii

Institute of Chemical Physics of the Academy of Sciences of the USSR, Vorobjevskoje Shausse, 2-b, Moscow, 117334, USSR

The similarities and differences between electron-nuclear tunnelling (discovered in 1966) and molecular tunnelling (discovered in 1973) are described here in terms of radiationless electron transitions. Examples of the low-temperature limit on a chemical reaction rate, observed since 1973, include: growth of chains of polymerisation of formaldehyde, Fe-CO rebinding in haem-containing proteins, isomerisation of radical pairs, hydrobromination of ethene, and photoinduced conversion of rhodopsin into prelumirhodopsin. The general significance of molecular tunnelling for chemistry and related subjects is discussed. The possible manifestations of molecular tunnelling in the formation of complex molecules at the surfaces and in the bulk of 'dirty ice' mantles of dust grains in interstellar dense clouds are also considered.

ONE of the most important consequences of the wave properties of matter is tunnelling—the ability of particles to penetrate the potential barriers whose heights exceed the particles' kinetic energy; the sub-barrier region is therefore classically forbidden for such particles.

Tunnelling starts to become important when the de Broglie wavelength of a particle becomes comparable to the barrier width in a way that the probability of tunnelling decreases with an increase of the width (d), the height (E) of the potential barrier, and the mass (m) of the tunnelling particle.

The tunnelling concept had its first successful application in nuclear phenomena such as α -decay^{1,2}, spontaneous fission, and thermonuclear reactions and later greatly contributed to solid state physics, electronics³⁻⁵ and more recently to cosmology⁶.

Chemical conversions

Such concepts in chemistry were initially concerned with the delocalisation of particles between two identical potential wells⁷, and led to analysing theoretically the tunnelling contribution to chemical reaction rates⁸⁻¹¹. Such contributions should be particularly significant at low temperatures because—in accordance with Arrhenius' law—the rate of classical transitions over the barrier decreases exponentially with decreasing temperature while the probability of simple sub-barrier tunnelling should level off as $T \rightarrow 0$, rather than drop towards zero.

The quite general criterion of a 'tunnelling temperature'² is given by $T_t = (\hbar/k_B \pi^{1/2} d)(E/m)^{1/2}$ where k_B is the Boltzmann constant. For $T < T_t$ tunnelling starts to dominate exponentially the Arrhenius-type transitions and the rate of exothermic reactions in many common problems gradually reaches its low-temperature plateau.

Generally an exothermic transition $i \rightarrow f$ requires the mixing of an initial level U_i^0 with final levels U_f^n and the dissipation of the transition energy ΔU_{if} provided by the fast electronic-vibrational relaxation of the final state (with the rate of such dissipation being larger than the rate of transition itself).

Different theoretical treatments lead either to the limit of the statistically averaged rate of such transitions or to their complete disappearance at low temperatures. Hence tunnelling does not necessarily produce the low-temperature plateau of chemical conversion rates, however, its occurrence is a very strong argument in favour of the quantum tunnelling mechanism of these conversions.

For several decades the search for chemical tunnelling was restricted to the region of comparatively large temperatures ($T > T_t$), and conclusions about its existence were based on minor (up to tens of per cent) isotope effects in the rate of conversion of hydrogen-containing and deuterated compounds¹³.

However, even stronger kinetic isotopic effects in such conversions (up to factors of $\exp(500/T)$) still do not confirm tunnelling because they can be caused simply by differences in zero-vibrational energies of Z-H and Z-D bonds.

Experimental observations of chemical tunnelling began in 1966 when Chance *et al.*^{14,15} discovered the low-temperature (120–4K) plateau (at $W \sim 300 \text{ s}^{-1}$) of the reaction rate of the laser-flash initiated oxidation of cytochrome *c* by chlorophyll. Using the simplest model of tunnelling as being a penetration of a plane wave through the one-dimensional barrier and taking the pre-exponential factor as $\omega_e \sim 4 \times 10^{15} \text{ s}^{-1}$ (Gamov factor $G_t = (W/\omega_e) \sim 10^{-13}$) and the activation energy $E \approx 0.14 \text{ eV}$ (as obtained from experiments up to 300 K) Chance *et al.*^{14,15} estimated the tunnelling length of an electron as $l \approx 30 \text{ \AA}$. A similar value could be obtained by comparing the position of the low-temperature plateau with the tunnelling temperature T_t defined above.

Calculated tunnelling lengths of such an order of magnitude are also typical of many other examples of low-temperature transfer of electrons—in biopolymers and in irradiated frozen glasslike solutions—if they are treated simply using Gamov factors and assuming that the barrier height equals the activation energy in a high-temperature region.

However, such tunnelling resembles an electron current

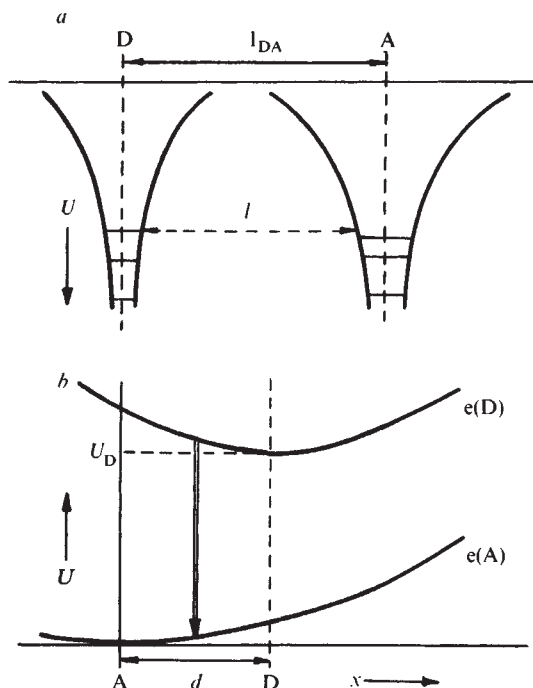


Fig. 1 Scheme of electron-nuclear tunnelling for electron transfer from the donor (D) to the acceptor (A). a, The tunnelling of electron over the distance l (levels in D and A do not usually coincide). b, The displacement of nuclei over the distance d for the transition of the electron $e(D) \rightarrow e(A)$ indicated by vertical arrow.

through the thin dielectric layer which breaks the electron circuit rather than the transfer of reactants in the chemical reaction.

The general situation is much more complex because chemical tunnelling of electrons is related to the displacement of nuclei and we need to consider possible violations of the Franck-Condon principle. This was noted for the first time by Libby¹⁶, and was later analysed in several studies¹⁷⁻²⁰ of the tunnelling of electrons in oxidation-reduction interactions of ions in solutions and in electrode reactions.

After Chance's discovery^{14,15} when it seemed necessary to consider the electronic-vibrational interactions (vibrational coupling), the Franck-Condon factors in the tunnelling-type oxidation-reduction transformation of complex bimolecules were based on ideas of radiationless electron transitions (see refs 21-25). Different descriptions of electron tunnelling in biopolymers as generalised electron-nuclear tunnelling are described in refs 26-30.

In 1973 the discovery of a low-temperature limit of the rate of complex chemical reactions was first reported^{31,32} using the example of the growth of polymerisation chains of formaldehyde, FA (see also refs 33-35).

We shall emphasise here that the words 'chemical reactions' describe transformations which include the reconstruction of molecules, the spatial rearrangement of atoms, changes in all main characteristics of valence bonds, their multiplicities and nature (σ or π -bonds) as well as lengths and angles.

The theoretical treatment of this phenomenon^{31,32} as a new kind of molecular chemical tunnelling used the Siebrand's theory³⁶ of radiationless relaxation of electronically-excited molecular states. Later this theory was also extended to the case of molecular tunnelling, the so-called tunnel-polaron non-Arrhenius chemical reaction kinetics³⁷.

Numerous surveys have been devoted to tunnelling of electrons in chemistry and biology (see refs 38-40) and we touch here only briefly the theory of such tunnelling to emphasise both the similarities and the differences of electron-nuclear and molecular tunnelling.

The probability of a radiationless electron transition W_{if} is determined by the product of the matrix element of the electron

transition $L_e^2 = |\langle \psi_{ei} | \hat{L} | \psi_{ef} \rangle|^2$ and of the Franck-Condon factor $F_v = |\langle \psi_{vi} | \psi_{vf} \rangle|^2$:

$$W_{if} = (2\pi/\hbar) L_e^2 F_v \rho_f = \omega_{eff} F_v$$

where ψ_e and ψ_v are the electron and vibrational wave functions respectively, $\hat{L}$ is a transition operator such as the non-adiabaticity operator and $\rho_f = (\hbar\omega_f)^{-1}$ is the density of vibrational levels in the final state.

The wave functions $\psi_e(x)$ are proportional to $\exp(-x/\alpha)$ for $x > \alpha$, where $\alpha \sim \hbar/(2mE)^{1/2}$ is the damping parameter, that is the size of the region of localisation of electron density, which is close to the length of chemical valence bond a (at $E \sim 1$ eV).

The wave functions $\psi_v(x)$ are proportional to $\exp[-(x/\Delta)^2]$ for $x > \Delta$, where $\Delta = (\hbar/M\omega)^{1/2}$ is the amplitude of nuclear vibrations; ω is the frequency of these vibrations; M the mass of the nuclei.

Among the most important parameters of electron and molecular tunnelling are the lengths of tunnelling (generally configurational displacement) given below as l for electrons and as d for heavy particles.

In most cases of electron-nuclear tunnelling in oxidation-reduction processes²⁶⁻³⁰: $l \gg \alpha \approx a \gg d \approx \Delta$.

The long-range transfer of electrons from the donor to the acceptor is not accompanied by chemical reaction. Displacements of donor and acceptor nuclei proceed in a similar way to ordinary intramolecular radiationless transitions up to distances of the order of $\Delta \sim 0.1$ Å (see Fig. 1).

In the case of molecular tunnelling in chemical reactions we have $d \approx l \approx a \gg \Delta$.

For sufficiently large electron ($l \gg \alpha$) and nuclear ($d \gg \Delta$) displacements, when the matrix elements of the transition probability are determined by the asymptotics of wave functions, the matrix element of the electron transition transforms into the ordinary Gamov-type tunnelling factor, that is $L_e^2 \propto \exp[-\psi l(mE)^{1/2}/\hbar]$ (m is the electron mass, $\psi \sim 1$ depends on the shape of the barrier), while the vibrational (nuclear) Franck-Condon factor depends exponentially on the square of the nuclear displacement: $F_v \propto \exp[-(d(M\hbar\omega)^{1/2}/\hbar)^2]$ (here $\omega = \omega_i \ll \omega_f$)³⁵.

Parameters of various processes interpreted in terms of radiationless electron transitions are summarised in Table 1.

Just as Chance's discovery^{14,15} is generally regarded as a first experimental proof of the existence of chemical electronic tunnelling, various examples of a low-temperature plateau of the rate of chemical reactions can be considered as experimental confirmation of the existence of molecular chemical tunnelling.

The first such example as mentioned above was the discovery of a limit (below ~ 12 K) of the rate of growth of chains in

Table 1 General characteristics of processes interpreted in terms of radiationless electron transitions

	l	ω_{eff}/ω_0^*	d	F_v	$W_{exp}(s^{-1})$
1. Electron-vibrational relaxation	$\sim \alpha$	$> 10^{-3}$	$\sim \Delta$	$> 10^{-3}$	$10^{14}-10^8$
2. Tunnelling of electrons in oxidation-reduction conversions:					
(i) not taking into account the nuclear displacement;	$(10-30)\alpha$	$< 10^{-8}$	0	1	10^6-10^{-2}
(ii) taking into account the electronic-vibrational coupling	$< 10\alpha$	$< 10^{-5}$	$\sim \Delta$	$> 10^{-3}$	
3. Molecular tunnelling in chemical reactions:					
(i) 'pure' molecular tunnelling;	$\sim \alpha$	$> 10^{-3}$	$\sim a$	$< 10^{-4}$	$10^{10}-10^{-4}$
(ii) tunnel-polaron transitions	$\sim \alpha$	$> 10^{-3}$	$d_i \sim \Delta$ $\sum d_i \sim a$	$> 10^{-3}$	

* $\omega_0 = 10^{14} s^{-1}$ is a typical value of pre-exponential factor in chemical reactions.

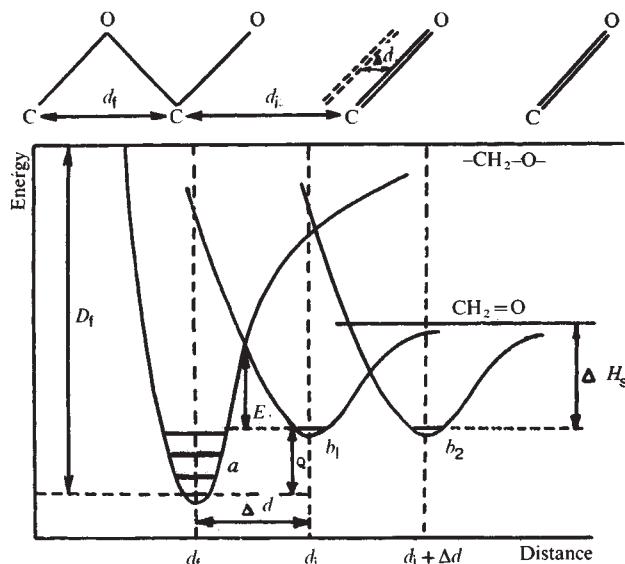
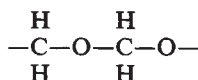


Fig. 2 The upper simplified scheme shows positions and displacements of molecules of solid formaldehyde and the polymer chain. The lower scheme illustrates the tunnel transitions of $\text{CH}_2\text{-O}$ groups during polymerisation ($b_1 \rightarrow a$) and during filling up of 'vacancies' formed near the end groups of polymer chain during polymerisation ($b_2 \rightarrow b_1$). The curve a is the adiabatic curve for a formaldehyde molecule added to a growing polymer chain; b_1 is a similar curve for the initial monomer molecule, and b_2 represents a similar curve for the next monomer molecule (the zero on the x -axis for this curve has been shifted by a distance Δd). The energy of dissociation of the C-O bond is $D \approx 4$ eV, energy of the sublimation of crystalline FA is $\Delta H_s \approx 0.3-0.4$ eV, polymerisation heat $Q \approx 0.37$ eV, activation energy $E = 0.1$ eV. Note that the product $\Delta d(M_{\text{CH}_2\text{O}})^{1/2}$ is close to the corresponding value for spontaneous fission of atomic nuclei, where $\Delta d \approx 10^{-12}$ cm and $E \sim 1$ MeV.

radiation-induced polymerisation of solid formaldehyde³¹⁻³⁵ studied by calorimetric method between 140 and 4 K.

Stacks of planar triangles of monomer CH_2O molecules with double C=O bonds transform in this reaction into long chains of single tetrahedral



bonds, and the substitution of van der Waals radii of C and O (typical for monomers) by much shorter valence $-\text{C}-\text{O}$ bonds in polymers leads to a considerable ($\sim 40\%$) increase in density. The scheme used for the calculation of the tunnelling rate in formaldehyde polymerisation is illustrated by Fig. 2.

In such an analysis the rate of growth of the polymer chain depends exponentially on the square of the molecular tunnelling distance $(\Delta d)^2$ of the monomer group. Besides the addition of new links to the growing chain, the polymerisation includes in this calculation the subsequent stage of delocalisation of the cavity formed at the border between the end link of the polymer chain and the neighbouring monomer molecule—provided, for example by quantum diffusion⁴¹⁻⁴³.

Another variant shown in Fig. 3 is the replacement of a well-defined cavity between the polymer and monomer by an extended distorted region of variable density (such as the stretched end of a polymer chain or the 'swollen' monomer that fills the cavity) which plays to some extent the part of polaron-deformed region of the crystal and is transferred (following the transfer of 'driving' particle such as an electron) with the end of the growing chain, corresponding to the above mentioned tunnel-polaron variant of non-Arrhenius kinetics³⁷ listed among the other mechanisms—in Table 1.

Adopting a complex (generally, multidimensional) picture of potential surfaces means we can consider the possibility of conversions through low barriers: in other words, the possibility of eliminating various restrictions on the tunnelling rate which

are in ordinary molecular tunnelling due to large masses or large displacements of reacting species.

The temperature dependence of the rate constant on the growth of polymerisation chains of formaldehyde is shown in Fig. 4.

The average time of adding one new link to the growing polymer is $\tau_0 = k^{-1} = \tau/G$ (where τ is the time taken for the growth of the whole chain and G the radiation-chemical yield of polymerisation) which is $\tau_0 \approx 10^{-2}$ s at the plateau.

Frauenfelder *et al.*⁴⁴⁻⁴⁵ confirmed the existence of molecular tunnelling in a study of low-temperature Fe-CO reconstitution in a haem-containing protein and protohaem. Fe-CO bonds were broken by lasers flashes and the kinetics of their reconstitution was spectrophotometrically followed over very wide intervals of time (10^{-6} – 10^3 s) and temperatures (2–340 K).

Besides finding a low-temperature plateau below ~ 10 K (see Fig. 4) these observations demonstrated the complex, polychromatic character of these kinetics, that is the existence of broad spectra of activation energies and barrier widths rather than single values of these parameters.

The polychromatic kinetics described in refs 48–49 and later combined with the tunnelling concepts^{50,51} is typical for systems with many different reactant states represented, for example, by different conformers of proteins or by traps of different depths in irradiated molecular solids.

Frauenfelder *et al.*⁴⁴⁻⁴⁷ successfully used polychromatic kinetics to solve the complex picture of several potential barriers hindering Fe-CO reconstitution and to correlate these barriers with the structures of protohaem, myoglobin and haemoglobin.

The low-temperature rate limit of the processes studied by Frauenfelder *et al.* was also demonstrated by Mössbauer absorption spectroscopy^{52,53}.

Meanwhile using Mössbauer emission spectroscopy (in particular, on combining it with the technique of delayed $\gamma\gamma$ -coincidences) allowed electron-nuclear tunnelling to be observed which caused the change of state of metallic ions in iron-cyanide⁵⁴ and cobalt-phenantrolyne⁵⁵ complexes.

A Japanese research group⁵⁶ has recently extended the earlier ESR observations⁵⁷ of conversions of radicals in γ -irradiated dimethylglyoxime at 77–220 K to lower temperatures (4–77 K) and succeeded in finding the low-temperature plateau of the rate of the comparatively slow process of isomerisation of radical pairs (see Fig. 4) which is characterised at higher temperatures

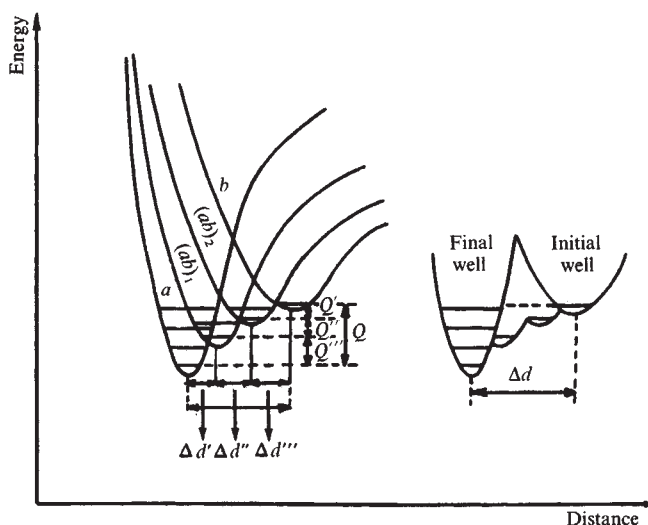


Fig. 3 The adiabatic potential curves for the normal (a) and 'stretched' (b) groups of polymer chains and CH_2O molecules in the crystal of a monomer. In the case of a well-defined hole the curves a and b represent two neighbouring states separated by a potential barrier (shown in the figure) which prevents elementary chain growth. In the case of a 'stretched' polymer chain there are two possible intermediate states: $(ab)_1$ and $(ab)_2$ between a and b . Any barriers are absent in transitions of the type $b \rightarrow (ab)_2 \rightarrow (ab)_1 \rightarrow a$.

by an activation energy $E \approx 0.65$ eV. Having studied the kinetics of hydrobromination for an equimolar ethene-HBr mixture between 90 and 30 K, our group observed long ($\nu \sim 200$) chains of this reaction even^{34,58} at 30 K. Subsequent calorimetric investigations of stationary and non-stationary kinetics of chain hydrobromination in the conditions of predominantly quadratic or linear rupture of chains led⁹⁰ to the numerical determination of the rate constant of propagation of chains ($k_p = \tau_0^{-1} = G/\tau$) and to finding another example of the limit of the rate of elementary chain reactions at low temperatures (see Fig. 4).

Application of picosecond laser spectroscopy to the studies of photo-initiated rhodopsin-prelumirhodopsin conversion⁵⁹ has demonstrated once more a low-temperature reaction rate plateau.

However, there is no agreement about the mechanism of this process: there exist two main approaches, namely the tunnelling of one or two hydrogen atoms to the nitrogen or Schiff base (across a distance of 0.5–0.9 Å)⁵⁹ or the low-temperature *cis-trans* isomerisation of the chromophore group (retinal) bound to the protein (opsin)^{60–62}.

The large reaction rate at the plateau ($K_H \approx 2.8 \times 10^{10} \text{ s}^{-1}$) and a quite weak kinetic isotopic effect ($K_H/K_D \approx 7$)—see Fig. 4—remind us of the possibility of a tunnel-polaron mechanism³⁷ of low-temperature isomerisation—via several states represented by intermediate potential curves in Fig. 3. Confirmation of this idea would be highly significant to the problems of low-temperature quantum chemistry, therefore, it is very important to make a definite choice between the transfer of hydrogen atoms in rhodopsin and the isomerisation of retinal.

The tunnelling of atoms and molecules with a rearrangement of chemical bonds should not be treated as exclusive to low-temperature chemistry. Such tunnelling can also be observed as sub-barrier predissociation of molecules by very precise laser tuning of their excitation energy below the Landau-Zener gap (this possibility was mentioned in ref. 38) or by electron

beams⁶³. Isotopic differences in the probabilities of such tunnelling may be useful for isotope enrichment.

One of the new topics of macroscopic chemical kinetics which could result from tunnelling is the cold initiation of thermal explosions by high pressures which enhance the tunnelling as well as the possibility of oscillations between subcritical and overcritical states of the system compared with the conditions of thermal explosion.

It is also possible that in some cases the usual transition from the diffusion to the kinetic region with decreasing temperature will be supplemented by a return to the diffusion region below the tunnelling temperature—because of the gradual disappearance of temperature dependence on the rate of the kinetic stage of conversion.

The high selectivity of low-temperature tunnelling reactions caused by the interplay of the three main parameters— d , M and E may be advantageous for applied cryochemistry and cryobiology.

The persistence of chemical reactivity even near absolute zero makes it doubtful whether a complex organism could be brought back to life after being subjected to prolonged freezing in conditions when chemical chain reactions might be initiated by external agents like radiation.

Of special interest is the possibility of chemical and prebiotic evolution at very low temperatures when exothermic reactions (including the formation of highly complex products) are thermodynamically favoured (because all entropy restrictions are eliminated as $T \rightarrow 0$) and at the same time—in spite of the Arrhenius law—quantum effects permit measurable, finite rates of chemical conversions.

The problem of cold tunnelling evolution from the 'storage of negative entropy' up to prebiotic levels has been discussed previously^{30–34,64–66}. The possibility of tunnel-polaron-type³⁷ low-temperature exothermic reactions of various kinds (such as polymerisation and polycondensation) could be particularly significant.

Interstellar chemistry

In view of all these circumstances we now consider the role of low-temperature quantum effects of chemical kinetics in the formation of complex products of interstellar chemistry. However, we should emphasise that—in contrast with the experimental facts and theoretical considerations described above—the following comments are purely hypothetical and they are intended to stimulate discussions and new experiments. The problems of interstellar chemistry have been treated in many recent surveys (see refs 67–77, 91), and we will consider them only to the extent needed for discussing the possibilities of molecular tunnelling in interstellar chemical reactions.

The most abundant element in the Universe is hydrogen. The cosmic abundance of helium (relative to hydrogen) is 0.09, followed by oxygen (7×10^{-4}), carbon (3×10^{-4}), nitrogen (9×10^{-5}), then magnesium, silicon and iron ($\sim 10^{-4}$ altogether). Amongst almost 50 interstellar molecules detected by radio and optical spectroscopy the most widespread apart from H_2 are: CO ($\sim 10^{-4}$), NH_3 , HCN and HNC ($\sim 10^{-6}$). There is no doubt that water and methane are also well represented in interstellar space, and we should also mention formaldehyde ($\sim 10^{-8}$) whose low-temperature polymerisation was directly observed in the laboratory.

The main stage of interstellar chemistry is the multitude of so-called interstellar clouds—giant (10^{17} – 5×10^{19} cm) suspensions of dust grains in a gas medium. The total dust amounts to only about 1% of the gaseous mass; dust grains are generally assumed to consist of central 'cores' ($r \geq 10^{-6}$ cm) of silicates or graphite surrounded by mantles of 'dirty ice' ($r \sim 10^{-5}$ cm) formed by a mixture of frozen substances. The average number of hydrogen atoms per dust grain in the cloud is 10^{12} .

One of the main differences between diffuse ($n_H \sim 10$ – 10^3 cm^{-3}) and dense or dark ($n_{H,H_2} \sim 10^3$ – 10^7 cm^{-3}) clouds is the following: diffuse clouds are transparent to UV radiation which causes ionisation of atoms and dissociation of molecules, while

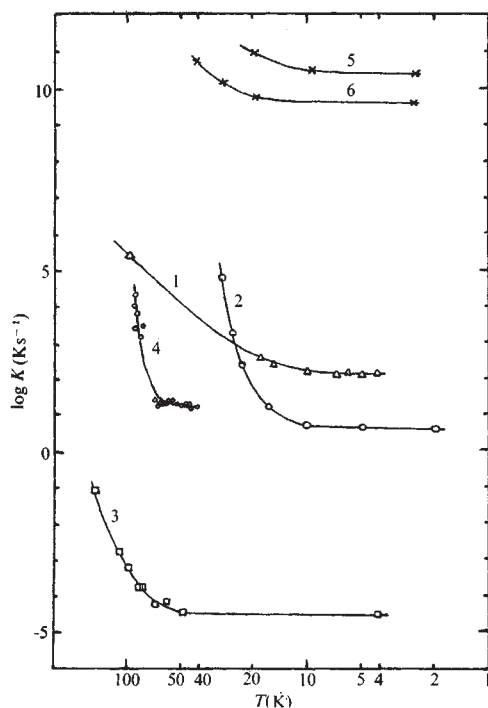


Fig. 4 Summary of data on the low-temperature limits of rate constants (K s^{-1}) of chemical reactions ($\log K$ – $\log T$ coordinates). Curve 1 (Δ), growth of chains of polymerisation of formaldehyde. Curve 2 ($\circ$), reconstitution of Fe–CO bonds in β -haemoglobin. Curve 3 ($\square$), isomerisation of radical pairs in γ -irradiated dimethylglyoxime. Curve 4 ($\circ$), propagation of chains in the hydrobromination of ethene. Curve 5 ($\times$), photoinduced conversion of rhodopsin into prelumirhodopsin. Curve 6 ($\times$), photoinduced conversion of deuterated rhodopsin into prelumirhodopsin.

dense clouds are opaque to UV radiation. Therefore photodissociation does not proceed in the depths of dark clouds and a much wider variety of complex, m -atomic molecules (up to $m = 11$, cyanotetra-acetylene HC_5N) are present here. This is why dark clouds are also referred to as molecular clouds. The temperature of dust grains in these clouds is $T = 10\text{--}20\text{ K}$.

The outer regions of dark clouds are transparent to UV radiation (its flux $\mathcal{F}_{\text{UV}} \sim 3 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ and its cross-section for interaction with molecules is $\sigma_{\text{UV}} \sim 10^{-18} \text{ cm}^2$), and each molecule absorbs a UV quantum within $\tau_{\text{UV}} \sim 100 \text{ yr}$.

The only external agents which can ionise or excite molecules inside the clouds are energetic cosmic protons. They interact with each interstellar molecule within $\tau_p \sim 3 \times 10^9 \text{ yr}$. This value strongly exceeds not only τ_{UV} , but also the lifetime of the clouds themselves which has been determined by their collisions and by gravitational collapse to be $\tau_{\text{cl}} \sim 10^5\text{--}10^7 \text{ yr}$.

In discussing interstellar chemistry we need to consider three types of chemical reactions—in the gas phase, at the surfaces of dust grains, and in the bulk of dirty ice. Tunnelling plays no part in gas phase reactions and they will not be considered here.

As far as reactions at the surfaces of dust grains are concerned, most attention has been paid to the activationless recombination of free atoms and radicals^{69,70,75}. According to existing estimates⁷⁰ tunnelling can play an important and even decisive part in lateral diffusion processes which precede such recombination for physically and chemically adsorbed hydrogen atoms and for heavier atoms and radicals in the case of their physical adsorption.

What this means to molecular tunnelling (or tunnel-polaron transitions) as a possible mechanism of chemical reactions at the surface and in the bulk of dust grains, has been discussed recently^{64,65}. An analysis of the possible role of tunnelling in such conversions should take into account the competition between the chemical reactions of complex molecules (components of dirty ice), their ejection from surfaces into the gas and their destruction by interstellar radiation which can not only decompose more complex molecules into simpler species but also initiate exothermic chain reactions of molecular combination for example, polymerisation. The destruction effects of radiation seem to restrict the formation of complex, m -atomic molecules in gas-phase reactions and in the recombination of radicals at grain surfaces to $m \sim 10\text{--}12$. Therefore such reactions cannot account for the formation of more complex interstellar molecules—up to polymer size.

Meanwhile astrophysicists such as Sagan^{73,76}, Greenberg^{72,77,78}, Wickramasinghe^{79,80} and Hoyle^{81–83}, although differing in their conjectures and conclusions about the nature of very complex interstellar molecules and their origins, nonetheless agree that such molecules do exist in the interstellar dust. These conclusions are based both on laboratory analyses of products of low temperature UV illumination of solid mixtures resembling the dirty ice^{73,77,78} (however, these analyses were performed after defrosting such mixtures, and therefore indicated the products of not only the 'cold' reactions but also of their after effects) and on the comparison of galactic absorption and emission IR spectra with the laboratory spectra of trial compounds^{79–83,76}.

Sagan considers the tar-like macromolecular compounds of irregular composition and structure formed in his laboratory model experiments as likely interstellar compounds and called⁷⁶ them tholins.

After comparing various terrestrial and interstellar IR spectra Hoyle and Wickramasinghe arrived at some interesting although controversial conclusions about the formation of interstellar polymers of formaldehyde, of polyoxymethylene^{79,80} and even of polysaccharides^{81–83}.

All these authors agree that the formation of any polymer requires the growth of macromolecules in a chain reaction, but they differ in the astrophysical conditions deemed necessary for such chain reactions.

According to Sagan^{73,76}, polymers and other sufficiently thermostable complex organic substances were transported into

the interstellar gas as components of grains ejected from solar nebulae by radiation pressure and stellar winds at the early stages of stellar evolution. Furthermore, tens of identified interstellar m -atomic molecules ($m = 2\text{--}11$) were formed not in the usual processes of integration in the gas phase and at the surface of grains but as degradation products of more complex interstellar dust compounds. Although we shall not discuss this point of view here, it should be noted that polymers above a certain temperature T_d are quite often not truly stable but only metastable, when stabilised by endcapping groups (see ref. 84). Being formed below T_d such polymers remain stable even above T_d , but polymer chains will not grow at $T > T_d$.

Another idea is that the polymerisation (as well as the formation of complex molecules by recombination of radicals^{69,70,72,75,77,78}) proceeds directly in the interstellar dust grains at $T = 10\text{--}20\text{ K}$, in the bulk of the ice mantles or at their surfaces. Even in the depths of dense clouds polymerisation chains can be triggered by long-range cosmic protons ($E_p \geq 100 \text{ MeV}$). However, before the discovery of the low-temperature limit of the chemical reaction rate speculations about the hypothetical formation of polymers within dense clouds lacked kinetic grounds.

At present we can attempt quite a detailed analysis of such possibilities for at least one example, the polymerisation of formaldehyde.

Interstellar polymerisation

Let us start with the abundance of formaldehyde in dirty ice. The saturated gas phase concentration of CO in equilibrium with solid carbon monoxide at 20 K is $\sim 10^5 \text{ cm}^{-3}$, and for methane $\sim 1 \text{ cm}^{-3}$. These values are larger or fairly close to the real concentrations of CO and CH_4 in the interstellar gas which, in such circumstances, can be undersaturated. Therefore the freezing-out of CH_4 (and especially of CO) seems to have much less significance on the formation of dirty ice than the solidification of ammonia, HCN, HNC, water and ethane.

Characterising the volatility by the temperature which corresponds to an equilibrium gas concentration of 1 cm^{-3} , then we have in order of decreasing volatility: ethane ($\sim 35\text{ K}$), ammonia and formaldehyde ($\sim 60\text{ K}$), HCN and HNC, then water ($\sim 90\text{ K}$) and, finally, a polymer of formaldehyde-polyoxymethylene ($\sim 100\text{ K}$). The rate of condensation does not depend on volatility while the rate of sublimation (both spontaneous and radiation-induced) decreases rapidly with diminishing volatility. Therefore these consequences accord with the steady-state enrichment of dirty ice by various components in comparison with gas. While the relative abundance of formaldehyde in the gas phase is no larger than several tenths of a per cent of all the other listed compounds, its abundance in the solid phase, in its polymer form, could be much higher. Moreover, we should consider that the above listed cosmic abundances of various elements provide only for water and formaldehyde the possibility to form the whole mass of dust grains (0.01 of hydrogen mass) while the maximum total mass of interstellar ethane is half of that, of HCN and of HNC a quarter of that, and of ammonia a sixth of that.

The attainment of steady-state conditions requires perceptible gas-solid transitions in both directions, that is a combination of condensation and sublimation. The condensation rate is high enough: the number of molecules of a certain kind condensed at each grain during the cloud life-time τ_{cl} (years) is $100 \kappa n \tau_{\text{cl}} (T/M)^{1/2}$, where M is their molecular mass (in daltons); κn (cm^{-3}) their gas phase concentration ($n = n_{\text{H}_2}$, $\kappa \leq 10^{-6}$); $\varepsilon \leq 1$ their sticking probability. For the most dense and long-lived clouds ($n \tau_{\text{cl}} \geq 10^{12}$) the condensation can even lead to a several-fold increase of the masses of grains during the time that the cloud exists.

Estimates of the sublimation rate contain more uncertainties. The rate constant of spontaneous (thermal) evaporation of molecules bound to the surface with a binding energy D equals $K_{\text{se}} \sim 10^{12} \exp(-D/k_B T)$, and the time of such an evaporation of the surface monolayer would exceed the clouds

lifetime if $(D/k_B) \geq 60 T$.

Detailed estimates of the rate of induced evaporation caused by the transient heating of the entire grain by low-energy cosmic rays and, to a lesser extent, by the spot heating of the same origin are given in ref. 70.

According to these estimates the lifetime of clouds ensures numerous repetitions of evaporation of molecules physically adsorbed at the strange lining (at $D \approx 0.05\text{--}0.1$ eV). However, the evaporation of mantle would be strongly hindered by its conversion to a crystal with greater binding energy. There will be no evaporation at all of molecules incorporated into the polymer chain. Therefore one can expect formaldehyde to be more abundant in the surface layers of interstellar grains because of its polymerisation.

The radiation-chemical yield of solid-state polymerisation of crystalline formaldehyde at 20 K is $G \sim 10^4$ molecules per 100 eV of radiation energy³¹⁻³⁵. Thus the energy delivered by cosmic protons to a grain during $\sim 3 \times 10^5$ yr is already sufficient to cause its complete polymerisation if it were to consist exclusively of formaldehyde: mixed composition of dirty ice certainly hinders polymerisation. Nevertheless, although the active end group is surrounded from all sides but one by admixtures it will 'automatically' select the only possible direction to add the neighbouring monomer molecule. The formation of a new link of the growing chain at the surface can be preceded by several futile encounters of the end group with other molecules hitting the surface from gas phase or diffusing along the surface. It would be interesting for interstellar chemistry to compare the G values for low-temperature polymerisation initiated by electrons and γ rays (when the local heating along the tracks is insignificant), by protons, and by heavy ions (when the heating along the tracks is particularly strong⁸⁵). It would also be desirable to determine directly the radiation yield of the evaporation of various molecules from different surfaces.

The rate of lateral diffusion of physically adsorbed molecules at low temperatures is determined by tunnelling and an estimate⁷⁰ of the time of a quantum jump of a molecule with $M = 30$ gives $t_{tr} \sim 0.1\text{--}1$ s. For the diffusion in the 'strange matrix' when the encounter of reactants requires N jumps, the reaction rate would be determined either by diffusion (at $\tau_0 < t_{tr}$), or by kinetic factors (at $\tau_0 > t_{tr}$), and the time of reaction would correspondingly be the time of N or of $(N\tau_0/t_{tr})$ jumps. If the end group of a growing chain is surrounded only by its own monomer, the diffusion factors cease to play any part, and the possibility or impossibility of effective polymer accumulation in the interstellar dust grains is determined by the interplay of four characteristic times— τ_0 , τ_{uv} (for diffuse clouds and outer layers of dense clouds), and τ_{cl} (deep in dense clouds), as $\tau_{cl} < \tau_p$.

Consideration of these criteria leads to the following inequalities required for the realisation of chain reactions with chain length ν , in dense clouds:

$$\tau_0 \ll (10^{-3}\text{--}10^{-2})\tau_{uv}/\nu \quad \text{or} \quad \tau_0 \ll \tau_{cl}/\nu$$

Taking the value of the pre-exponential factor $\omega_0 \approx 10^{14} \text{ s}^{-1}$ one gets the following condition for the classical (Arrhenius-type) interstellar chain reaction to proceed: $E < T(5.6 \times 10^{-3} - 2 \times 10^{-4} \log \nu)$ eV.

In the case of polymerisation of formaldehyde ($\nu \geq 10^3$, $E_{\text{exp}} \approx 0.1$ eV, refs 31-35) the inequality $\tau_0 \ll 10^2\text{--}10^4$ yr would be fulfilled for Arrhenius kinetics only at $T > 20$ K, while the experimental value is $\tau_0 \approx 10^{-2}$ s, at $T = 4\text{--}12$ K.

An approximate representation of the rate of tunnelling is obtained by using Gamov factors (see refs 37, 66): $G_t = \exp(-\beta d(mE)^{1/2}/\hbar)$. This yields the following condition for quantum (tunnelling) chain polymerisation of formaldehyde in dense clouds:

$$\beta d(mE)^{1/2}/\hbar < 4 - 0.15 \log \nu$$

where $\beta = 2\sqrt{2}$ for a rectangular barrier, $\beta = \pi\sqrt{2}/2$ for a parabolic barrier, and $\beta = 4\sqrt{2}/3$ for a triangular barrier, d is in angstroms, and E in electron volts. This inequality corresponds to: $d(E)^{1/2} \leq 60\text{--}90$ for the tunnelling of electrons, $d(E)^{1/2} \leq 1.4\text{--}2$ for H atoms, $d(E)^{1/2} \leq 0.35\text{--}0.5$ for formaldehyde mole-

cules in a monomer matrix.

In this way tunnelling could significantly increase the number of possible low-temperature reactions in dense clouds. One case would be, for example, the possibility of tunnelling polymerisation at the surface of dust grains with the formation of a very thin (several molecular layers) polymer film around the inner region of dirty ice. Such a film could protect the surface of this inner region from both condensation and sublimation.

For chemical and prebiotic evolution the reactions of polycondensation in dirty ice mantles with the participation of CH_2O , HCN, HNC, NH_3 and H_2O are of interest. Such reactions lead to the formation of amino acids, polypeptides, sugars and nucleotide bases (purines and pyrimidines), they are exothermic, but not chain-type.

There are no reasons why there could be a 'pure' molecular tunnelling mechanism of such reactions—the rate of tunnelling falls steeply to a vanishingly small limit with increasing barrier widths and masses. However, each single step of a chemical conversion which represents an elementary gas phase process (such as the reaction $\text{H}_2\text{C}=\text{O} + \text{NH}_3 \rightarrow \text{H}_2\text{C}=\text{NH} + \text{H}_2\text{O}$), proceeds in the solid as a sequence of many individual or collective conformational rearrangements of molecules, complexes or indeed whole regions of molecular crystals. The collision of a dust grain with a cosmic proton or UV quantum, or the release of recombination energy at the grain surface can induce the transfer of the 'driving' particle, such as the electron, which determines the number of conformational rearrangements. In this way the tunnel-polaron mechanism³⁷ becomes involved: the single long-range transfer of a molecule ($d \approx a$) is replaced by several short-range displacements ($d_i \approx \Delta$, $\sum d_i \approx a$) which can lead to the lowering or even to the disappearance of the activation barrier (see Fig. 3), and to a reasonably high and temperature-independent rate of chemical conversions of cold systems such as of isomerisation. This mechanism could be important also for chemical and prebiotic evolution.

One should, however, bear in mind that if a certain product is formed as a final step in a non-chain sequence of η chemical conversions, and each of these conversions needs an external stimulus, then the yield of such product would be $\eta\nu$ times smaller than the number of molecules entering the chain of length ν . Therefore the yield of multistep polycondensation reactions accumulated during the life-time of dense clouds would not exceed 0.01% of the mass of the grains.

The formation of even the most complex biopolymers in interstellar dust grains does not guarantee their conservation during the later formation of new stars and planetary systems, when the dense clouds collapse. One can not exclude, for example, the strong heating at the surfaces of planets at certain stages in their evolution. Nevertheless, one should also keep in mind the variety of possible chemical reactants and reactions in 'warm' systems depending on their history. For example, polymers created and stabilised by endcapping at low temperatures could survive the consequent warming, but they cannot be formed directly at higher temperatures. Meanwhile the existence of such polymers opens additional possibilities of chemical conversions not only in cold, but also in warm systems, such as the integration of molecules in shock-wave induced solid-phase reactions^{86,87} and in the vicinity of phase transitions, which proceed during the alternate heating and cooling of reactants^{88,89}. Favourable conditions for the latter type would be provided, for example, by multiple transportation of stable organic compounds between circumstellar disks and interstellar clouds⁷⁶ and in comets with extended orbits which alternately suffer short heating when approaching the Sun, and prolonged deep cooling, away from the Sun.

The hypothesis of a cold 'entropyless' exothermic formation of complex molecules permitted only by quantum chemical reactivity is not an alternative, but an additional idea on the mechanisms of chemical and prebiotic evolution in terrestrial and extraterrestrial conditions.

I thank Professor V. A. Benderskii for stimulating discussions.

1. Gamov, G. Z. *Phys.* **51**, 204 (1928).
2. Gurney, R. W. & Condon, E. U. *Nature* **122**, 439 (1928).
3. Esaki, L. *Science* **183**, 1149 (1974).
4. Giaever, I. *Science* **183**, 1253 (1974).
5. Josephson, B. D. *Science* **184**, 527 (1974).
6. Hawking, S. W. *Scient. Am.* **236**, 34 (1977).
7. Hund, F. Z. *Phys.* **43**, 805 (1927).
8. Bourgin, D. G. *Proc. Natn. Acad. Sci. U.S.A.* **15**, 357 (1929).
9. Roginsky, S. Z. & Rozenkevitch, L. Z. *phys. Chem.* **B 10**, 47 (1930).
10. Wigner, E. P. Z. *phys. Chem.* **B 19**, 1903 (1932).
11. Bell, R. P. *Proc. R. Soc.* **139**, 466 (1933); **148**, 241 (1935); **158** (1937).
12. Goldanskii, V. I. *Dokl. Akad. Nauk SSSR* **124**, 1261 (1959); **127**, 1037 (1959).
13. Bell, R. P. *The Proton in Chemistry* 2nd edn (Cornell University Press, Ithaca, 1973).
14. De Vault, R. & Chance, B. *Biophys. J.* **6**, 825 (1966).
15. Chance, B., De Vault, D., Legallais, V., Mela, V. & Yonetani, T. in *Fast Reactions and Primary Processes in Chemical Kinetics* (ed. Claesson, S.) 437 (Firth Nobel Symp., Stockholm, Almqvist a. Wiksell, 1967).
16. Libby, W. F. *J. phys. Chem.* **56**, 863 (1952).
17. Marcus, R. J., Zwolinski, B. J. & Eyring, H. *J. phys. Chem.* **58**, 432 (1954).
18. Marcus, R. J. *J. chem. Phys.* **24**, 966 (1956).
19. Levich, V. G. & Dogonadze, R. R. *Dokl. Akad. Nauk SSSR* **124**, 123 (1959).
20. Levich, V. G. *Adv. Electrochem. Engng.* **4**, 249 (1965).
21. Fong, F. K. (ed.) *Radiationless Processes in Molecules and Crystals* (Springer, Berlin, 1976).
22. Robinson, G. W. in *Excited States* (ed. Lim, E. C.) 1 (Academic, New York, 1974).
23. Freed, K. F. *Radiationless Processes in Polyatomic Molecules—in Stereo—and Theoretical Chemistry*, 105 (Springer, Berlin, 1972).
24. Jortner, J. in *2nd Int. Symp. Organic Solid-state Chemistry*, 389 (London, Butterworth, 1971).
25. Henry, B. R. & Kasha, M. A. *Rev. Phys. Chem.* **19**, 161 (1968).
26. Grigorov, L. N. & Chernavskii, D. S. *Biofizika* **17**, 195 (1972).
27. Blumenfeld, L. A. & Chernavskii, D. S. *J. theor. Biol.* **39**, 1 (1973).
28. Hopfield, J. J. *Proc. Natn. Acad. Sci. U.S.A.* **71**, 3640 (1974).
29. Hopfield, J. J. *Biophys. J.* **18**, 311 (1977).
30. Jortner, J. *J. chem. Phys.* **64**, 4860 (1976).
31. Goldanskii, V. I., Frank-Kamenetskii, M. D. & Barkalov, I. M. *Dokl. Akad. Nauk SSSR* **211**, 133 (1973).
32. Goldanskii, V. I., Frank-Kamenetskii, M. D. & Barkalov, I. M. *Science* **182**, 1344 (1973).
33. Goldanskii, V. I. *Russ. Chem. Rev.* **44**, 1019 (1975).
34. Goldanskii, V. I. A. *Rev. Phys. Chem.* **27**, 85 (1976).
35. Goldanskii, V. I. in *Tunneling in Biological Systems* (Academic, New York, 1977).
36. Siebrand, W. J. *chem. Phys.* **46**, 440 (1967); **47**, 2411 (1967).
37. Goldanskii, V. I. *Doklady Akad. Nauk SSSR* **239**, 104 (1978).
38. Libby, W. F. A. *Rev. Phys. Chem.* **28**, 105 (1977).
39. Zamarayev, K. I. & Khairutdinov, R. F. *Chem. Phys.* **4**, 181 (1974).
40. Goldanskii, V. I., Zamarayev, K. I., Mikhailov, A. I. & Khairutdinov, R. F. in *Kinetics of Elementary Chemical Reactions* 68 (Nauka, Moscow, 1973).
41. Andreev, A. F. & Lifshits, I. M. *Soviet Phys. JETP* **29**, 1107 (1969).
42. Kagan, Yu. & Maksimov, L. A. *Soviet Phys. JETP* **38**, 307 (1973).
43. Kagan, Yu. & Klinger, M. I. *J. Phys. C* **7**, 2791 (1974).
44. Alberding, N. et al. *Science* **192**, 1002 (1976).
45. Alberding, N. et al. *J. Chem. Phys.* **65**, 4701 (1976).
46. Eisenstein, L. *Int. J. Quantum Chem.: Quantum Biol. Symp.* **21** (1976).
47. Frauenfelder, H. in *Tunneling in Biological systems*, University of Pennsylvania (1977).
48. Primack, W. *Phys. Rev.* **100**, 1677 (1955).
49. Primack, W. *J. appl. Phys.* **31**, 1525 (1960).
50. Mikhailov, A. I., Kuzina, S. I., Lukovnikov, A. F. & Goldanskii, V. I. *Dokl. Akad. Nauk SSSR* **204**, 383 (1972).
51. Mikhailov, A. I., Bolshakov, A. I., Lebedev, Ya. S. & Goldanskii, V. I. *Fizika tverd. Tela*, **14**, 1172 (1972).
52. Marcolin, H. B., Reschke, P. & Trautwein, A. in *Proc. int. Conf. Mössbauer Spectr.* (eds Barb. D. & Tarina, D.) 301 (Bucharest, 1977).
53. Maeda, T., Harami, T., Sakai, H. & Morita, Y. in *Proc. int. Conf. Mössbauer Spectr.* (eds Barb. D. & Tarina, D.) 299 (Bucharest, 1977).
54. Alekseev, V. P., Goldanskii, V. I., Prussakov, V. E., Nefedjev, A. V. & Stukan, R. A. *Soviet Phys. JETP Lett.* **16**, 43 (1972).
55. Grimm, R., Gütlisch, P., Kankelait, E. & Link, R. *J. chem. Phys.* **67**, 5491 (1977).
56. Toriyama, K., Nunome, K. & Iwasaki, M. *J. Am. chem. Soc.* **99**, 5823 (1977).
57. Yakimchenko, O. E. & Lebedev, Ya. S. *Int. J. Radiat. Phys. Chem.* **3**, 17 (1971).
58. Kiryukhin, D. P., Barkalov, I. M. & Goldanskii, V. I. *Dokl. Akad. Nauk SSSR* **238**, 388 (1978).
59. Peters, K., Applebury, M. L. & Rentzepis, P. M. *Proc. Natn. Acad. Sci. U.S.A.* **74**, 3139 (1977).
60. Green, B. H., Monger, T. G., Alfano, R. R., Aton, B. & Callender, R. H. *Nature* **269**, 179 (1977).
61. Lewis, A. *Proc. Natn. Acad. Sci. U.S.A.* **75**, 549 (1978).
62. Warshel, H. B. & Deakyne, C. *Chem. Phys. Lett.* **55**, 459 (1978).
63. Erman, P. *Phys. Scripta* **16**, 60 (1977).
64. Goldanskii, V. I. *Nature* **268**, 612 (1977); **269**, 583 (1977).
65. Goldanskii, V. I. *Dokl. Akad. Nauk SSSR* **235**, 1053 (1977); **238**, 823 (1978).
66. Goldanskii, V. I. *Chem. Scripta* (in the press).
67. Gammon, R. H. *Chem. Engng. News* **56**, 21 (1978).
68. Watson, W. D. *Acct. Chem. Res.* **10**, 221 (1977).
69. Watson, W. D. *Rev. mod. Phys.* **48**, 513 (1976).
70. Watson, W. D. & Salpeter, E. E. *Astrophys. J.* **174**, 321 (1972).
71. Herbst, E. & Klempner, W. *Phys. Today* **29**, 32 (1976).
72. Greenberg, J. M. *Ned. Tidschr. Natuurk.* **42**, 117 (1976).
73. Sagan, G. *Nature* **238**, 77 (1972).
74. Rank, D. M., Townes, C. H. & Welch, W. J. *Science* **174**, 1083 (1971).
75. Millar, T. J. & Williams, D. A. *Mon. Not. R. astr. Soc.* **137**, 527 (1975).
76. Sagan, C. & Khare, B. N. *Nature* **277**, 102 (1979).
77. Greenberg, J. M. in *Molecules in the Galactic Environment* (eds Gordon, M. A. & Snyder, L. E.) (Wiley, New York, 1973).
78. Greenberg, J. M. *Astrophys. J. Lett.* **189**, 81 (1974).
79. Wickramasinghe, N. C. *Nature* **252**, 462 (1974).
80. Wickramasinghe, N. C. *Mon. Not. R. astr. Soc.* **170**, 11 P (1975).
81. Hoyle, F. & Wickramasinghe, N. C. *Nature* **268**, 610 (1977).
82. Hoyle, F. & Wickramasinghe, N. C. *Mon. Not. R. astr. Soc.* **181**, 51 P (1977).
83. Wickramasinghe, N. C., Hoyle, F., Brooks, J. & Shaw, G. *Nature* **269**, 674 (1977).
84. Fawcett, A. H. *Nature* **257**, 159 (1975).
85. Goldanskii, V. I. & Kagan, Yu. *Int. J. appl. Rad. Isot.* **11**, 1 (1961).
86. Barkalov, I. M. et al. *J. Polym. Sci. C* **2597** (1967).
87. Goldanskii, V. I., Ignatovich, T. N., Kosygin, M. Yu. & Yampolskii, P. A. *Dokl. Akad. Nauk SSSR* **207**, 218 (1972).
88. Kargin, V. A. & Kabanov, V. A. *Zhur. Vses. Khim. Obshch. im. D. I. Mendeleeva* **9**, 6026 (1964).
89. Kargin, V. A., Kabanov, V. A. & Papisov, I. M. *J. Polym. Sci.* **4**, 767 (1964).
90. Kiryukhin, D. P., Barkalov, I. U. & Goldanskii, V. I. *J. chim. Phys.* (in the press).
91. Anders, E., Hayatsu, R. & Studier, U. H. *Science* **182**, 781 (1973).

articles

New large-scale magnetic features of the Milky Way

M. Simard-Normandin & P. P. Kronberg

University of Toronto, Scarborough College and David Dunlap Observatory

A large body of newly determined rotation measures (RM) for extragalactic radio sources gives a new, much clearer picture of the magnetic structure of our Galaxy. The RM sky is dominated by a very strong zone which extends to large (down to $b \approx -40^\circ$) negative galactic latitudes. Both the strength and extent of this feature are greater than appreciated from earlier data. It coincides in angular position with loop II of the non-thermal radio background radiation of the Galaxy. This, and two other strong RM features suggest that our Galaxy has very large-scale magnetic field zones, probably between the spiral arms, and that the prevailing sense of the field is opposite in the interarm regions on each side of the Sagittarius arm.

WE report the discovery of some new large-scale features of the galactic magnetic field which have been obtained from a programme to determine accurate rotation measures (RM) of a large sample of extragalactic sources.

Previous analyses of Faraday rotation¹⁻⁷ have shown that a systematic modulation of the RMs occurs with galactic coordinates (l, b) and that RMs are systematically larger at low galactic latitudes. They also indicate a prevailing sense of RM consistent with a local magnetic field directed towards $l \approx 90^\circ$ (refs 2, 4, 6, 7), a trend which also agrees with that deduced from the RMs of pulsars⁸.

Our new polarisation data have greatly improved the number and quality of RM determinations for extragalactic radio sources, and currently give 528 RMs. This is nearly a factor of 2 more than the number of previously published RMs of equivalent reliability. The increased density of sources on the

galactic (l, b) sky provides us with the clearest picture yet obtained of the large-scale galactic contribution to the RMs.

A large, high-latitude anomaly in the RM sky

Figure 1 shows the distribution in galactic coordinates of rotation measures of 476 extragalactic sources. This sample has been purged of 52 sources whose RMs, although accurately determined, are likely to be 'contaminated' by Faraday rotation in the sources themselves, rather than in the interstellar medium. We have established that sources with high depolarisation rates (that is, low $\lambda_{1/2}$ values; see ref. 9) have systematically higher RMs. We have therefore removed sources with low $\lambda_{1/2}$ (<10 cm), as this is an independent source parameter unlikely to be related to the effects of our Galaxy.

The most prominent feature in the RM sky is a large zone of negative RMs from $b = +10^\circ$ to -40° at $45^\circ \leq l \leq 160^\circ$. We shall denote this as region A. Although previous RM data showed the region centred near $l \sim 90^\circ$ to have predominantly negative RMs, the large angular scale and strength of this region are revealed for the first time by our new data. Figure 1 shows that surprisingly large RMs, up to 200 rad m^{-2} , occur at negative latitudes as far as -30° near $l \sim 90^\circ$. This contrasts with most other galactic regions, in which the RMs are small at $|b| \geq 12^\circ$.

Other new zones of high RM off the galactic plane

To display the large-scale RM features better, we have calculated the average RM of neighbouring sources within a 10° radius of each source. Sources with RMs greater than 1.3σ from the mean are omitted from the averaging calculation. Figure 2 shows the average RM centred on each source. Inspection of Fig. 2 reveals that there are two other zones of systematically large RM above $|b| = 5^\circ$; one (region B), centred near $l = 255^\circ$, extends to negative latitudes $b \sim -20^\circ$, and has large positive RMs of up to 300 rad m^{-2} . A third high-RM zone (region C) at positive latitudes is centred at $l \sim 40^\circ$, $b \sim +5^\circ$.

We note that in the south galactic hemisphere the prevailing RM sense as previously noted by Gardner, Morris and

Whiteoak², and Vallée and P.P.K.⁵, is negative in $0^\circ \leq l \leq 180^\circ$ and positive at $180^\circ \leq l \leq 360^\circ$. This trend continues from the galactic plane as far as the south galactic pole, and is reinforced by the present, larger RM sample. The overall trend of RMs in the northern galactic hemisphere (NGH) is less straightforward. At latitudes about $+20^\circ$ the prevailing RM sense is positive for $0^\circ \leq l \leq 90^\circ$ (which we denote as the first quadrant), and the sign alternates in successive quadrants in the NGH. One explanation of this pattern, which was also recognised by Gardner *et al.*² and Vallée and P.P.K.⁵, is that loop I (the North Galactic Spur, see ref. 10) seems to cause a perturbation of the local magnetic field on a large angular scale in the first and fourth quadrants ($0^\circ \rightarrow 90^\circ$, and $270^\circ \rightarrow 360^\circ$) in the NGH. The outer two quadrants reflect the trend of RM present at southern latitudes.

Rotation measures and the galactic loops

The perturbation of the RMs in the first and fourth quadrants and $b > +20^\circ$ is consistent with a large magnetic loop associated with loop I which is the largest 'off-plane' feature of the galactic continuum radiation¹¹. We can now establish the magnitude of this perturbation as $\approx 30 \text{ rad m}^{-2}$. This is much weaker than features A, B and C mentioned above. Feature C, being close to the plane, is probably not associated with loop I.

Loop III (ref. 12) in the northern hemisphere has no corresponding large-scale RM feature greater than $\sim 20 \text{ rad m}^{-2}$ (Figs 1, 2). The fact that loops I and III do not appear as strong features in the RM sky is consistent with other evidence that they are local features.

Loop II (ref. 12), on the other hand, seems to encircle RM feature A. Whether or not this is merely coincidence is not clear, but we also note that region A coincides with a mild enhancement of the galactic continuum radiation. If loop II is similar in nature to loops I and III and if it is also associated with feature A, the very large RMs in A require it to be too large to be local. A mean RM at $l = 90^\circ$, $b = -25^\circ$ of -150 rad m^{-2} implies, for $\langle n_e \rangle = 0.03 \text{ cm}^{-3}$ and an aligned magnetic field of $3 \times 10^{-6} \text{ G}$, a scale length of 2 kpc. We conclude that if loop II is a local feature associated with loop III (which is directly on the opposite side of the galactic plane), then it cannot be associated with RM feature A. We consider that loop II is unlikely to be a local feature in view of its close angular coincidence with feature A.

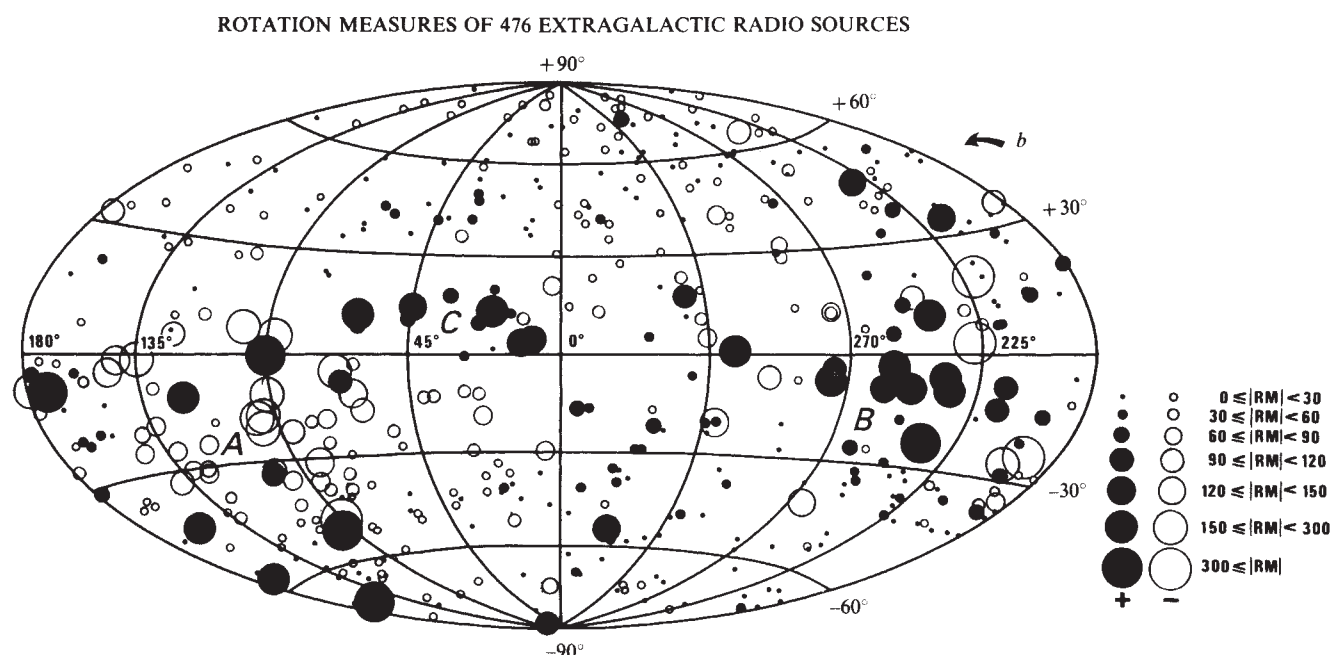


Fig. 1 RMs of 476 extragalactic radio sources displayed on an equal area projection. Features A, B and C (see text) are labelled.

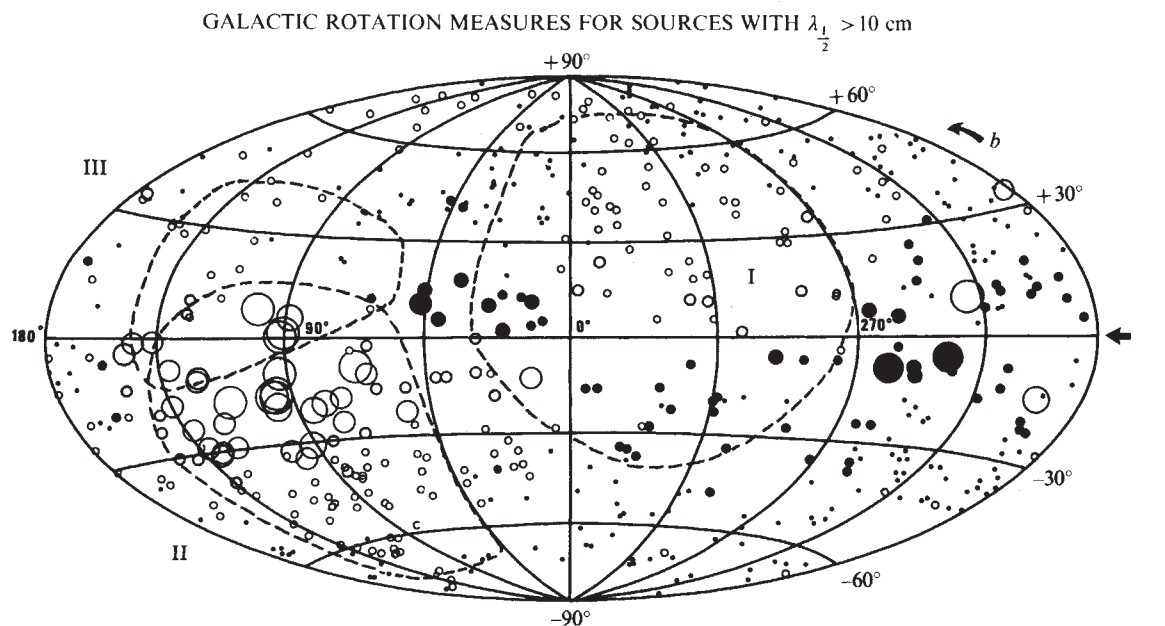


Fig. 2 Mean RMs in circles of 10° radius centred on each source, which emphasise the large-scale variations better than Fig. 1. The RM scale is the same as that of Fig. 1. The positions of the radio continuum loops I, II and III are outlined. Note that some small-scale features which are real are smoothed out, and the precise location of boundaries where the RM changes sign are slightly masked. These are best located in Fig. 1.

The large scale of feature A (implied by its large RM) requires that it have a large magnetic energy, $\gg 10^{52}$ erg. On this consideration alone, RM feature A is too energetic to be associated with a supernova remnant. On energetic grounds it could be associated with either galactic rotation or large-scale streaming of the infalling high-velocity HI clouds, of which there are several in the same direction^{13,14}. It is possible that feature A is associated with local streaming effects which are related to the magellanic stream, because the end of the magellanic stream lies in the approximate direction of feature A. More data are required, however, to establish any connection with the magellanic stream.

The structure of magnetic field

The large rotation measures observed near the galactic plane around $l \approx 255^\circ$ (feature B) and $l \approx 90^\circ$ are consistent with a longitudinal magnetic field directed towards $l \approx 90^\circ$. In the direction opposite to galactic rotation the large RM zone is centred near 255° rather than 270° which suggests that the Perseus arm may be opening up in this direction. The RMs and dispersion measures (DM) of pulsars in this direction indicate a mean magnetic field of less than $3 \mu\text{G}$ (90° and 270° south of the galactic plane). From this we conclude that the extent of the aligned field in both these areas (A and B) is greater than ~ 4 kpc. A and B must therefore be large-scale features of the Galaxy. Feature C at $l \approx 45^\circ$ can be interpreted as another major longitudinal component located approximately between the Sagittarius and the Norma-Scutum arms where (as the magnitude of the RM suggests) the prevailing field is directed along our line of sight for several kiloparsecs.

The very abrupt reversal of RM near the plane at $l \approx 60^\circ$ is consistent with a model containing two prevailing field zones pointing in opposite directions along the spiral interarms. This is best shown in Fig. 3, in which we superimpose the model of the spiral structure of the Galaxy by Georgelin¹⁵.

The fields do not seem to be co-planar. They also extend to quite large angular distances from the mean galactic plane. An interesting question at this stage is whether a counterpart to, say, region C can be seen on the $l < 360^\circ$ side of the Galaxy—in the southern hemisphere. Unfortunately, we have not made

measurements from the southern hemisphere, and have calculated RMs only from polarisation measurements in the literature for this region. It is now clearly of interest to obtain more data at $|b| < 30^\circ$ in the range $360^\circ \geq l \geq 260^\circ$. This would verify if large-scale magnetic field components can be seen associated with arm or interarm regions on the southern side of the Milky Way.

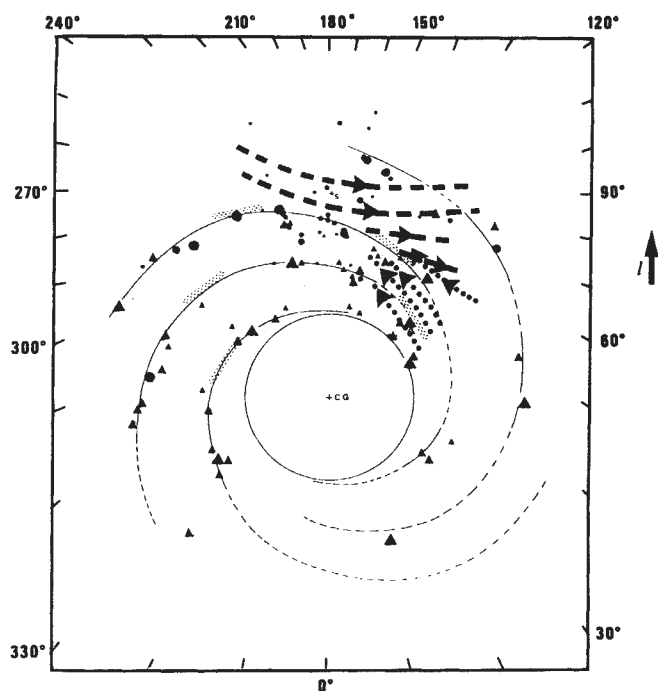


Fig. 3 A sketch of the regions containing an aligned component of magnetic field, indicating their approximate scale and direction. This is shown superimposed on the distribution of bright HII regions in the Galaxy by Georgelin¹⁵. Note that the heavy dotted and dashed lines do not necessarily represent the local prevailing magnetic field direction, but rather the zones in which the line-of-sight integral of RM is large and has the same sense.

RMs of pulsars in the direction $l = 60^\circ$ near the plane indicate a more turbulent field exhibiting several reversals along the line of sight. Interarm magnetic fields, on the other hand, tend to be more homogeneous, as suggested by the fact that there are virtually no small rotation measures in features *A* and *B*.

Conclusion

We conclude that galactic loops I and III cause at most weak perturbations in the RM sky. Loop II may or may not be a local loop like loops I and III, depending on whether or not it is

associated with RM feature *A*. The latter must be at least 2 kpc and possibly as large as 6 kpc in length.

Weak fields ($<3 \mu\text{G}$) seem to occupy the interarm regions and maintain the same prevailing sense for distances of several kiloparsecs. A reversal of the prevailing field direction occurs near $l = 60^\circ$. The field direction seems to change abruptly as our line of sight crosses the Sagittarius arm. A more detailed and quantitative analysis of the galactic RM and magnetic structure will be published separately.

We acknowledge financial support from the NRC of Canada in the form of a grant (P.P.K.) and a scholarship (M.S.-N.). We also thank Stuart Button for assistance with the data processing.

Received 19 December 1978; accepted 13 March 1979.

1. Gardner, F. F. & Davies, R. D. *Aust. J. Phys.* **19**, 129–139 (1966).
2. Gardner, F. F., Morris, D. & Whiteoak, J. B. *Aust. J. Phys.* **22**, 813–819 (1969).
3. Kawabata, K., Fujimoto, M., Sofue, Y. & Fukui, M. *Publ. astr. Soc. Jap.* **21**, 293–305 (1969).
4. Vallée, J. P. & Kronberg, P. P. *Nature phys. Sci.* **246**, 49–51 (1973).
5. Vallée, J. P. & Kronberg, P. P. *Astr. Astrophys.* **43**, 233–242 (1975).
6. Kronberg, P. P. & Simard-Normandin, M. *Nature* **263**, 653–656 (1976).
7. Wright, W. E., thesis, California Inst. Technol. (1973).
8. Manchester, R. N. *Astrophys. J.* **188**, 637–644 (1974).

9. Kronberg, P. P., Conway, R. G. & Gilbert, J. A. *Mon. Not. R. astr. Soc.* **156**, 275–282 (1972).
10. Large, M. I., Quigley, M. J. S., Haslam, C. G. T. *Mon. Not. R. astr. Soc.* **131**, 335–350 (1966).
11. Landecker, T. & Wielebinski, R. *Aust. J. Phys. Suppl.* **16** (1970).
12. Berkhuijsen, E. M., Haslam, C. G. T. & Salter, C. J. *Astr. Astrophys.* **14**, 252–262 (1971).
13. Verschuur, G. L. *Astr. Astrophys.* **22**, 139–151 (1973).
14. Hulsbosch, A. N. M. & Oort, J. H. *Astr. Astrophys.* **22**, 153–154 (1973).
15. Georgelin, Y. M., thesis, Univ. Marseille (1976).

Preliminary correlations between the Koobi Fora and Shungura Formations, East Africa

T. E. Cerling*

Department of Geology, University of California, Berkeley, California 94720

F. H. Brown

Department of Geology, University of Utah, Salt Lake City, Utah 84112

B. W. Cerling, G. H. Curtis and R. E. Drake

Department of Geology, University of California, Berkeley, California 94720

Major, minor and trace element analyses of glass from pumice and tuffs from the Koobi Fora and Shungura Formations show that the Chari, Karari and Tuff L are correlative; that the KBS Tuff and Tuff H2 are correlative; and that an unnamed tuff in the Koobi Fora Formation is correlative with Tuff H4 of the Shungura Formation.

TWO important early Pleistocene hominid-bearing formations, the Koobi Fora Formation and the Shungura Formation, are found in the Lake Turkana Basin in northern Kenya and in southern Ethiopia, respectively (Fig. 1). These formations have yielded several hundred hominid specimens^{1,2} and several thousand artefacts^{3,4}. Despite almost 10 years of work in the region, there are no firm stratigraphic correlations between the two areas. Lack of exposure in the intervening area precludes correlations based solely on field studies.

Rhyolite tuffs, mostly reworked, occur in both regions and have served as marker beds for mapping, and have been useful for K–Ar dating. Over 100 tuffs occur in the Shungura Formation, and 12 of these are extensive enough to subdivide that formation into 12 members. The thick stratigraphic section, good exposure and small lateral facies variations have resulted in a well-established sequence^{5,6}. The Koobi Fora Formation was deposited near the basin margin by different, smaller river systems that drained the Suregei Cuesta and the Chew Bahir

(Lake Stephanie) region⁷. Mapping of the Koobi Fora Formation has been hampered by poor exposure and rapid lateral facies changes. Independent dating and palaeontological studies in the two regions led to faunal correlations that, if acceptable, are very disconcerting^{8,9}. If the chronologies of both formations were correct^{5,10}, faunas of similar evolutionary stage are separated by 0.5–1 Myr. Subsequent dating of pumice from the KBS Tuff (1.8 Myr)¹¹ in the Koobi Fora Formation did not substantiate the original date of 2.6 Myr^{10,12}. This discrepancy has led to additional studies^{13,14} which disagree with the younger date.

Volcanic ash is an excellent correlative tool in Quaternary studies. Mineral suites, refractive indices of minerals and glass, and the chemical composition of glass have all been successfully used to correlate air-fall tuffs^{15,16}. Reworked ash deposits, however, have additional complications. Erosion of pre-existing ash deposits and addition of pumice and glass from these deposits, abrasion of distinctive glass mantles on primary minerals and dilution of primary minerals by detrital minerals render several techniques useless. Primary minerals in Koobi Fora Formation tuffs are sometimes unrecognisable because of abrasion; electron microprobe studies of individual shards show that some tuffs are made up of several populations of shards with differing chemical compositions¹⁷. Because differing proportions of these populations would make trace element analyses of bulk glass samples of dubious value, the present study concentrated on characterising pumice fragments from tuffs because electron microprobe studies showed that the glass in a single pumice is homogeneous.

* Present address: Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830.

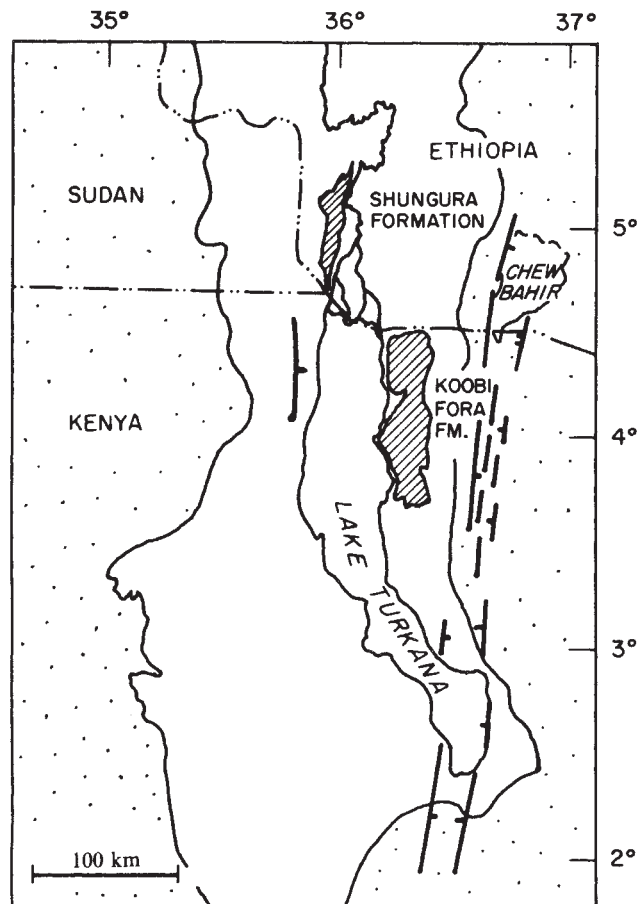


Fig. 1 Map of Koobi Fora and Shungura Formations.

Procedure

Tuff and pumice samples were collected during several field seasons (1972–77). Pure glass separates from individual pumice were obtained by gentle crushing and sieving, treatments with HCl and HF in an ultrasonic bath to remove adhering clay and silt particles, and repeated magnetic separations. Heavy liquids were used only as a last resort to keep bromine contamination to a minimum. Glass was mixed with cellulose and pressed into pellets for X-ray fluorescence analysis. USGS standards G-2 and GSP-1 were used, and results were corrected for absorption. Glass shards from tuff and pumice samples were analysed by electron microprobe for seven elements and were corrected for matrix effects¹⁸.

Results: tuff correlations and ages

Three tuffs stand out as having identical chemical properties and are thus considered to be the same tuff. Figure 2 summarises the chemical data for the Chari, Karari and KBS Tuffs from the Koobi Fora Formation and Tuffs H2, H4 and L from the Shungura Formation.

Chari–Karari–L: the Chari and Karari tuffs of the Koobi Fora Formation cap their respective sections at Ileret and along the Karari Ridge. These two tuffs have been thought to be correlative on the basis of stratigraphic position⁷, K–Ar dates¹⁰ and similar faunas in the underlying strata⁹. Trace and minor element contents of glass from pumice confirm this presumed correlation (Fig. 2). Electron microprobe analyses of individual shards from tuffs and of glass from pumice show that both have the same composition. One pumice cobble from Tuff L of the Shungura Formation has the same major, minor and trace element composition as pumice from the Chari and Karari Tuffs (Fig. 2, Tables 1, 2). A glass separate from Tuff L is nearly identical in trace element chemistry to that of the pumice (A. Mertz, personal communication). Feldspars from pumice in the

Chari and Karari Tuffs and from Tuff L range in composition from about $An_{10}Ab_{28}Or_{12}$ to $Ab_{64}Or_{36}$. Average calcium, iron and barium contents in these feldspars over the compositional range Or_{32} – Or_{36} are given in Table 3. The compositional range over which the average was taken is restricted because both barium and iron vary systematically with composition of the feldspar, and also to ensure comparability between different data sets on feldspars with different compositional ranges. K–Ar dating on nine feldspar separates range from 1.26 to 1.48 Myr and average about 1.35 Myr (R.E.D., in preparation). This agrees with previous results¹⁰.

As these tuffs have identical ages (R.E.D., in preparation) and chemical compositions, and because their feldspars are nearly identical in composition, they are probably products of the same eruption, and represent the same time horizon even though they were deposited by different drainage systems. This may indicate that the source area is on the drainage divide separating the two areas, as postulated earlier¹⁹.

Present nomenclature is confusing; the name 'Karari Tuff' should be eliminated as the name 'Chari Tuff' has precedence. Because both the terms 'Chari Tuff' and 'Tuff L' are firmly entrenched in the literature, however, we propose to use each term in its respective section.

KBS-H2: nine pumice fragments from the KBS Tuff in area 131, two pumice fragments from near the KBS type locality in area 105, and two pumice fragments from Tuff H2 were analysed (Fig. 2). Three distinctive pumice types are found in area 131: very dark brown glass with large vesicles containing about 5.5% Fe_2O_3 (one sample); grey glass with moderate vesicles having about 4.2% Fe_2O_3 (three samples); and silky white glass with stretched tubules having about 3.1% Fe_2O_3 (five samples). This tuff consists of platy and stretched glass

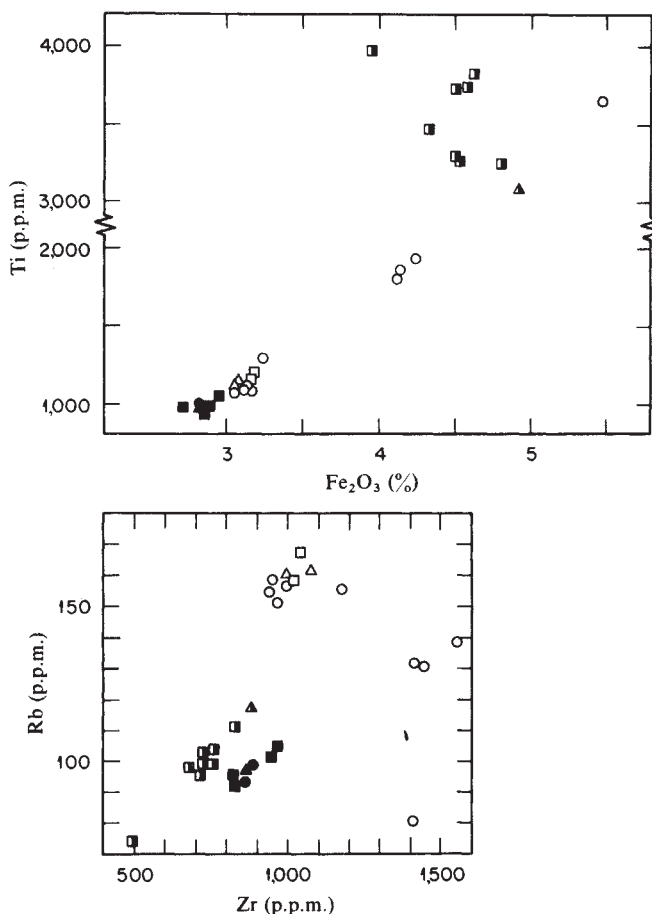


Fig. 2 Trace and minor element contents of pure glass separates from pumice from tuffs in the Koobi Fora and Shungura Formations. ●, Chari Tuff; ■, Karari Tuff; ▲, Tuff L; □, KBS Tuff-105-East; △, Tuff H4; ○, KBS-131; ◇, KBS-105; ▽, Tuff H2.

Table 1 Major and minor element compositions of glass from tuff and pumice

Unit	Material	Area	Sample no.	SiO ₂	Al ₂ O ₃	CaO	MgO	Fe ₂ O ₃	TiO ₂	MnO
Chari	Tuff	6	77-23	74.7	10.7	0.18	0.02	2.82	0.16	0.06
Chari	Pumice	6	6-001	75.8	10.5	0.17	0.02	2.76	0.16	0.05
Karari	Tuff	131	77-19	74.5	10.6	0.19	0.03	2.86	0.19	0.06
Karari	Pumice	118	108-001	75.4	10.5	0.19	0.02	2.86	0.17	0.05
Tuff L*	Pumice	Omo	76-L	74.6	10.5	0.19	<0.05	2.82	0.17	0.07
KBS	Tuff	102	72-44	73.9	10.9	0.19	0.05	3.07	0.18	0.11
KBS	Tuff	105	77-15	72.9	10.8	0.17	0.04	3.06	0.18	0.08
KBS	Tuff	131	75-2t	74.8	10.9	0.18	0.06	3.08	0.18	0.09
KBS	Pumice	131	75-2E	74.2	10.9	0.20	0.06	3.01	0.22	0.09
KBS	Pumice	131	75-2F	73.1	10.9	0.21	0.07	3.22	0.20	0.09
H2	Pumice	Omo	76-H2A	73.1	10.9	0.18	0.05	3.16	0.19	0.10
H2	Pumice	Omo	76-H2B	73.4	10.9	0.19	0.05	3.14	0.19	0.10
KBS	Pumice	131	75-2C	70.4	11.2	0.35	0.07	4.26	0.32	0.15
KBS	Pumice	131	75-2G	72.3	11.2	0.34	0.07	4.29	0.33	0.14
KBS	Pumice	131	131-002	72.0	10.9	0.34	0.08	4.38	0.32	0.14
KBS	Pumice	131	131-001	68.8	12.7	0.77	0.27	5.70	0.58	0.34
KBS*	Pumice	105-E	77-50A	70.6	10.6	0.34	0.3	4.80	0.54	0.24
H4*	Pumice	Omo	76-H4B	71.9	10.8	0.35	0.2	4.92	0.51	0.23

Determinations are by electron microprobe unless otherwise specified. Total iron as Fe₂O₃.

*Analysis by X-ray fluorescence.

shards; generally about 30% of these are brown to dark brown and close examination shows that there is a continuum from the colourless to the dark shards. Locally, basal portions of the tuff lack the dark shards. Electron microprobe studies show no difference in chemical composition between light and dark shards and indicate that they both have the same composition as the most common pumice type (that with about 3.1% Fe₂O₃, see Table 1). The three types are easily distinguished on the basis of their trace element chemistry (Table 2, Fig. 2). K-Ar dates were determined on all three types of pumice to determine whether older pumice had been reworked into a younger tuff. If this is the case, the pumices involved must differ in age by less than the analytical error (about ± 0.03 Myr), as the groups had average

dates of 1.82, 1.83, and 1.84 Myr, respectively (R.E.D., in preparation). Concordant glass-feldspar pairs were found for two pumice (R.E.D., in preparation).

Two pumice from the type KBS Tuff in area 105 have the same composition as the dominant group in area 131; electron microprobe analyses on glass shards from the tuff show that it is the same tuff as in area 131 (Table 1). A conventional K-Ar date of 1.75 Myr was obtained for one feldspar separate (R.E.D., in preparation). Two pumices from Tuff H2 in the Shungura Formation also have the same composition as the dominant pumice type in area 131 (Fig. 2). Electron microprobe analyses of glass from the pumice confirm this conclusion (Table 1). Four conventional K-Ar analyses (three from samples not analysed

Table 2 Results of X-ray fluorescence analyses on glass separates from pumice fragments

Tuff	Area	Sample no.	Nb	Zr	Y	Sr	Rb	Zn	Mn	Ti	Fe ₂ O ₃
Chari	6	6-001A	106	885	84	4	109	178	579	1000	2.82
Chari	1	1-9902CC	103	864	82	1	105	179	588	992	2.84
Karari	131	108-1	104	844	82	40	105	176	599	952	2.85
Karari	131	108-2	114	940	89	36	111	189	617	1059	2.95
Karari	131	108-3	107	832	81	9	103	178	573	980	2.71
Karari	131	108-4	118	964	92	5	115	183	595	996	2.88
Tuff L	Omo	76-L	105	867	90	7	107	176	576	999	2.82
KBS	131	75-2A	224	972	105	—	162	218	826	1129	3.13
KBS	131	75-2B	221	946	105	—	165	213	832	1080	3.05
KBS	131	75-2D	225	968	109	—	169	221	854	1098	3.16
KBS	131	75-2E	198	1002	115	11	168	219	841	1130	3.13
KBS	131	75-2F	213	1176	117	6	166	238	933	1305	3.24
KBS	105	77-108A	201	1016	113	4	169	220	836	1170	3.16
KBS	105	77-108B	219	1057	117	1	177	226	853	1203	3.17
H2	Omo	76-H2A	218	1014	110	4	170	220	835	1120	3.05
H2	Omo	76-H2B	201	1083	113	7	172	221	833	1146	3.06
KBS	131	75-2C	254	1412	109	—	142	252	1309	1811	4.11
KBS	131	75-2G	231	1436	110	5	141	254	1287	1889	4.14
KBS	131	131-002	239	1553	116	2	149	260	1373	1953	4.24
KBS	131	131-001	176	1408	76	5	91	198	2786	3611	5.48
KBS	105-E	77-109B	104	492	49	5	73	132	1615	3994	3.94
KBS	105-E	LUCAS-C	130	686	63	8	98	164	1762	3475	4.32
KBS	105-E	LUCAS-B	132	707	68	8	98	172	1877	3753	4.52
KBS	105-E	LUCAS-A	135	727	69	8	100	170	1847	3744	4.57
KBS	105-E	77-109D	154	741	67	15	103	154	1705	3293	4.51
KBS	105-E	77-109A	142	757	71	10	103	159	1767	3288	4.53
KBS	105-E	105-004	149	734	67	2	99	166	1882	3823	4.61
KBS	105-E	77-50A	171	830	74	19	111	189	1799	3245	4.80
H4	Omo	76-H4B	176	876	79	16	117	201	1822	3081	4.92

In p.p.m. except total iron, which is reported as % Fe₂O₃.

Table 3 Minor element content of feldspars

Sample no.	KA no.	Tuff	Area	CaO	BaO	Fe ₂ O ₃ *
108-1	2815	Karari	118	0.08	0.25	0.41
6-001A	2865	Chari	6	0.11	0.26	0.41
108-4	2880	Karari	118	0.09	0.31	0.36
108-2	2882	Karari	118	0.09	0.31	0.34
76-L	3254	Tuff L	Omo	0.10	0.29	0.36
75-2A	3160	KBS	131	0.08	0.07	0.43
75-2D	3161	KBS	131	0.06	0.06	0.43
77-108A	3189	KBS	105	0.08	0.05	0.38
76-H2A	3256	H2	Omo	0.08	0.06	0.42
75-2C	3166	KBS	131	0.09	0.06	0.36
75-2G	3167	KBS	131	0.08	0.02	0.45
131-002	2842	KBS	131	0.07	0.05	0.47
131-001	2869	KBS	131	0.28	0.11	0.38
77-50A	3182	KBS	105-E	0.14	0.03	0.65
	—	H4	Omo	0.10	0.04	0.69†

Averages taken over the range Or₃₂–Or₃₇.*Total iron expressed as Fe₂O₃.

†Analysis taken from ref. 5.

by X-ray fluorescence and from an earlier study⁹) average 1.87 Myr (R.E.D., in preparation).

There are several types of feldspars from the KBS Tuff. Those from pumice samples, where the iron content of the glass is about 3.1 or 4.2%, are indistinguishable chemically, and range from about An₈Ab₇₆Or₁₆ to Ab₆₂Or₃₈. Those from samples, where the iron content of the glass is about 5.5%, range from about An₁₆Ab₇₃Or₁₀ to Ab₆₃Or₃₇ and are distinctly higher in calcium and barium. Averages for all points with greater than 32% Or are given in Table 3. These feldspars are clearly distinct from the feldspars of the Chari-L Tuff. A third feldspar type is found in the KBS Tuff from area 105-E; it is distinct from the other two and is discussed below.

Discussion

The combined evidence justifies a correlation between the KBS Tuff of the Koobi Fora Formation in areas 131 and 105 and Tuff H2 of the Shungura Formation because of the close correspondence in age (R.E.D., in preparation) and composition. The age of 1.8 Myr for this tuff (R.E.D., in preparation) is compatible with palaeomagnetic studies²⁰, which place this tuff in a normally polarised section, and with faunal evidence⁹. We propose that this ash should continue to be called by its original term to avoid confusion in future studies.

KBS (area 105-East)-H4: the tuff that occurs in area 105-East is not the KBS Tuff because no pumice characteristic of that found in the KBS Tuff in areas 105 and 131 are found in this area. Compositions of glasses from pumice show that they are members of an alkaline-rhyolite differentiation series (Fig. 2, Table 2). A single pumice lump from Tuff H4 of the Shungura

Formation also falls on this differentiation trend. Composition of minor element chemistry of feldspars from pumice also show them to be related (Table 3, samples 77-50A and H4). Conventional K–Ar dates on these samples average 1.8 Myr. Although this tuff is similar in age to the KBS Tuff and Tuff H2, and also occurs in normally polarised sediments, stratigraphic relationships in the Shungura Formation show that Tuff H4 occurs about 15 m above Tuff H2²¹; this tuff is therefore younger than Tuff H2. Because these pumice from the KBS Tuff in area 105-East and Tuff H4 have indistinguishable chemistries, feldspar compositions, and conventional K–Ar dates (R.E.D., in preparation) they are probably the same tuff.

Correlation of the Chari and Karari Tuffs with Tuff L, of the type KBS Tuff with Tuff H2, and of the tuff in area 105-East with Tuff H4 of the Shungura Formation is independent of the controversy surrounding the age of the KBS Tuff^{9–14}. However, the correlations presented here are compatible with palaeomagnetic studies^{20,21} and additional K–Ar dates on other tuffs within the Koobi Fora, Kubi Algi and Shungura Formations⁵.

We thank Richard Leakey and the National Museum of Kenya for field support, and G. L. Isaac, R. L. Hay, A. J. W. Gleadow, J. W. K. Harris, I. C. Findlater, and A. M. Sarna-Wojcicki for helpful comments. Laboratory work was funded by the NSF.

Received 27 December 1978; accepted 12 March 1979.

- Day, M. H., Leakey, R. E. F., Walker, A. C. & Wood, B. A. *Am. J. Phys. Anthr.* **42**, 461–476 (1974).
- Day, M. H., Leakey, R. E. F., Walker, A. C. & Wood, B. A. *Am. J. Phys. Anthr.* **45**, 369–409 (1976).
- Harris, J. W. K. & Isaac, G. L. *Nature* **252**, 102–107 (1976).
- Isaac, G. L., Harris, J. W. K. & Crader, D. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. & Leakey, R. E. F.) 533–551 (University of Chicago Press, 1976).
- Brown, F. H. & Nash, W. P. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. & Leakey, R. E. F.) 50–63 (University of Chicago Press, 1976).
- de Heinzelin, J., Haesaerts, P. & Howell, F. C. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. & Leakey, R. E. F.) 24–49 (University of Chicago Press, 1976).
- Vondra, C. F. & Bowen, B. E. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. & Leakey, R. E. F.) 79–93 (University of Chicago Press, 1976).
- Maglio, V. J. *Nature* **239**, 379–385 (1972).
- White, T. D. & Harris, J. M. *Science* **198**, 13–21 (1977).
- Fitch, F. J. & Miller, J. A. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. & Leakey, R. E. F.) 123–147 (University of Chicago Press, 1976).
- Curtis, G. H., Drake, R. E., Cerling, T. E., Cerling, B. W. & Hampel, J. *Nature* **258**, 395–398 (1975).
- Fitch, F. J. & Miller, J. A. *Nature* **226**, 226–228 (1970).
- Hurford, A. J., Gleadow, A. J. W. & Naeser, C. W. *Nature* **263**, 738–740 (1976).
- Fitch, F. J., Hooker, P. J. & Miller, J. A. *Nature* **263**, 740–744 (1976).
- Izett, G. A., Wilcox, R. E., Powers, H. A. & Desborough, G. A. *Quat. Res.* **1**, 121–132 (1970).
- Sarna-Wojcicki, A. M. *US geol. Survey, Prof. Pap.* 972, (1970).
- Cerling, T. E. thesis, Univ. California, Berkeley (1977).
- Bence, A. E. & Albee, A. L. *J. Geol.* **76**, 382–403 (1968).
- Brown, F. H. in *Calibration of Hominid Evolution* (eds Bishop, W. W. & Miller, J. A.) 273–287 (Scottish Academic, Edinburgh, 1972).
- Hillhouse, J. W., Ndombi, J. W. M., Cox, A. & Brock, A. *Nature* **265**, 411–415 (1977).
- Brown, F. H. & Shuey, R. T. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. & Leakey, R. E. F.) 64–78 (University of Chicago Press, 1976).

Application of 'molecular' theories to the structure of the crystalline state

Jeremy K. Burdett

Chemistry Department, University of Chicago, Chicago, Illinois 60637

Recent years have seen the development of simple molecular orbital based ideas to view the structures of molecular ring, cage and cluster compounds. The principles of these methods are used here to demonstrate in a new way how the structures of solids depend on electronic configuration.

In contrast to the structures of simple molecules there are many possibilities for the geometrical arrangement of simple systems as extended arrays in solids. We may pinpoint three general factors which determine a solid state structure. First, the efficient packing together of the atoms in the solid. The occurrence of a particular AX structure will often be restricted if the radius ratio

of the two atomic species A and X is unfavourable. Second, the preservation of long-range order. Many structures look very favourable when a small cluster of atoms is built up but sometimes completion of another coordination shell is not possible whilst retaining the symmetry of the structure. Third, electronic forces. These may be divided into ionic, non-directional effects and covalent forces, directional in character. This article shows how recent developments in understanding the structures of molecular ring, cage, and cluster species may be used to provide a new way of viewing the structures of solids.

Theories for molecular species

One way, due to Mingos¹, that we can look at the geometries of these cages and rings is to start by taking the molecular orbital structure of a simple symmetric system and inquire what happens to the number of skeletal bonds as we add electrons to the system. Figure 1 shows a molecular orbital diagram for a tetrahedral A_4 species. It is characterised by the fact that there are six low energy orbitals holding the framework together and to higher energy six orbitals which are antibonding between the A atoms. The four orbitals which point outwards from the tetrahedron and do not take part in framework bonding are not shown. With five electrons per A atom there are just enough to fill all the six skeletal bonding orbitals (number of skeletal electron pairs, $n = 6$) and the outward pointing orbitals (lone pairs). This is the case for the electron precise tetrahedral molecule P_4 (Fig. 1) where there are six P-P linkages. For $n = 7$ two electrons must go into an antibonding orbital and one of the AA bonds is broken. Thus the structure found for B_4H_{10} is an opened-out tetrahedron. With $n = 8$ a similar structure results for $P_4(cy)_4$ (cy = cyclohexyl) but with no central linkage. For $n = 9$ three of the tetrahedral bonds must be broken. This is a novel way of looking at PCl_3 or S_2F_2 . The presence of a central atom does not disturb this simple picture (but we must be careful to count the number of skeletal electrons properly) and with 12 'skeletal' pairs CCl_4 has no bonds between the A (in this case Cl) atoms. An analogous scheme holds for the rationalisation of the structures of many transition metal cluster complexes which are difficult to understand in more conventional ways.

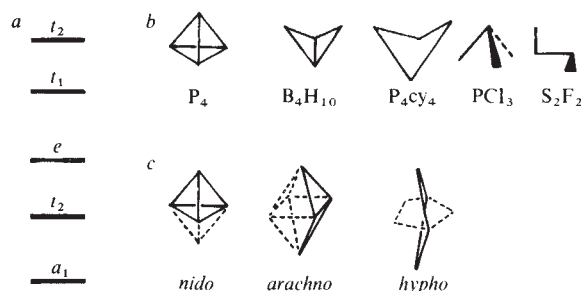


Fig. 1 a, Schematic molecular orbital picture of a tetrahedral A_4 species; b, the structures of some four atom units; c, the latter viewed using Wade's scheme.

The second approach, due to Wade^{2,3}, is again based on the number of skeletal electrons in the molecule, n . The resultant structure is based on that of an N vertex deltahedron where $N = n - 1$. If the carborane or cage compound has N skeletal atoms then all the vertices will be occupied (a *closo* structure) if $N - 1$ atoms all the vertices but one will be occupied (a *nido* structure), if $N - 2$ atoms then all the vertices but two will be occupied (an *arachno* structure), if $N - 3$ atoms then all the vertices but three will be occupied (a *hypho* structure), and so on. Figure 1 also shows how the P_4 , B_4H_{10} and P_4cy_4 molecules are viewed on this scheme, as *nido*, *arachno* and *hypho* species. Various rules⁴ have been developed to predict which vertices will be left vacant in *nido*, *arachno* and *hypho* structures.

Bond breaking and the number of valence electrons

A large number of AX systems with eight valence electrons per formula unit adopt the zincblende (sphalerite) or wurtzite structure (Fig. 2.) ZnS, carbon (diamond), AlSb, CuCl, CuBr, HgS are among the systems which adopt the former structure and

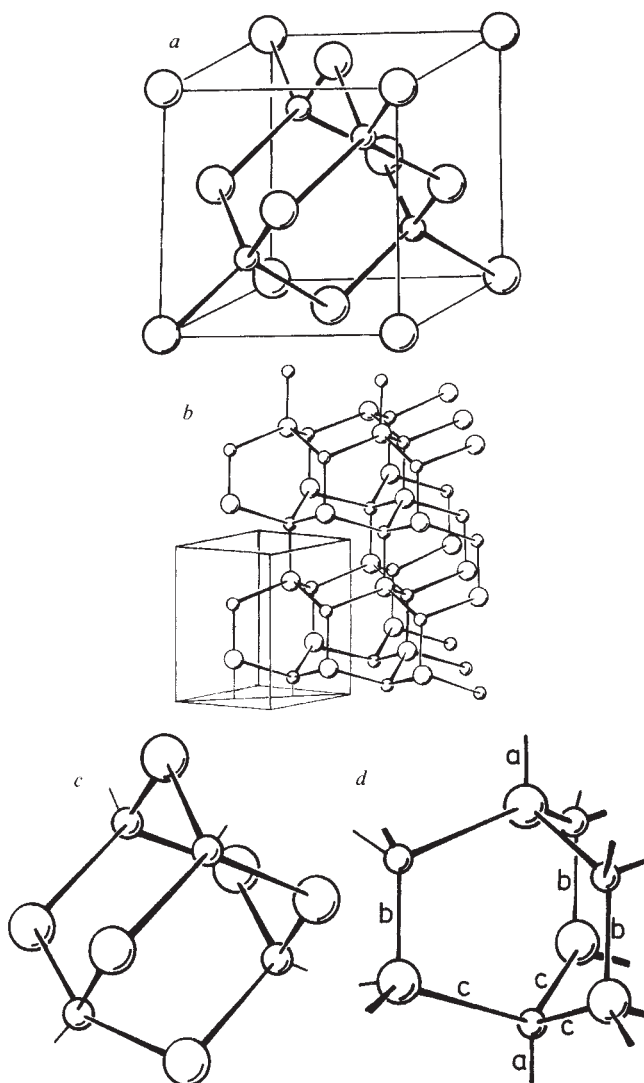


Fig. 2 a, The zincblende (sphalerite) and b, wurtzite structures; c, part of the zincblende structure indicating a 'unit cage'; d, unit cage of the wurtzite structure (other ways of choosing the cage are, of course, possible).

ZnO, AlN, BeO, CdS, CuI, MgTe the latter. Some of the examples are dimorphic. (For details of these and other systems cited here see refs 5-7.) Extraction of a chunk of lattice from each structure reveals the basic building block and demonstrates the similarity between the two systems. We recognise in these two lattice fragments the framework found for adamantane or P_4O_6 and P_4O_{10} (zincblende) or bicyclo(222)octane or barrellene (wurtzite) in molecular systems. How the molecule (or solid) is held together is best illustrated for the case of the diamond configuration. Each carbon atom has four, equivalent, tetrahedrally directed orbitals on a localised equivalent molecular orbital picture, made by suitable combination of 2s and 2p orbitals. With four electrons, each carbon atom contributes half a bonding electron pair to either a cage bond (that is towards the framework of the structures of Fig. 2c, d) or to an outward pointing bond which connects this cage to another. Thus no electrons go into antibonding orbitals and there are the right

number of electrons to make four bonds around each carbon atom. In the case of molecular P_4O_6 the outward pointing orbitals are filled with lone pairs of electrons leaving just 24 electrons, (that is 12 pairs) to make all 12 intracage bonds. These are examples of electron precise structures and regular geometries are found. But what happens when more than eight electrons per atom pair are present? According to the approach above, bonds are broken. Equation 1 shows one series of structures we will follow which show that the 'molecular' approach also works well for solids.

Wurtzite:	ZnS	$\rightarrow \beta$ GaSe	$\rightarrow$ As	$\rightarrow$ Se, Te	$\rightarrow$ I ₂	$\rightarrow$ Xe	1
No. of electrons:	8	9	10	12	14	16	

We start by considering the wurtzite structure. Each cage (Fig. 2d) contains two formula units. The three AX units forming the sides of the prism are shared by three such cages, the 'apical' AX unit is associated with only one cage. The apical atoms are attached to another cage by a single linkage top and bottom. The prismatic atoms are each attached to two other cages. There are thus 16 electrons per cage, all residing in bonding orbitals. With one extra electron per formula unit (for example GaSe) there are sufficient electrons to break one two-electron-bond per cage. We need to count up the number and type of the bonds present in the cage very carefully in order to see how many bonds may be broken. We may break either two extra-cage linkages, (a) in Fig. 2d, three prismatic cage linkages (b) or two pyramidal (c) linkages with these electrons. If extra-cage bonds are broken

then each cage will contribute to the bond breaking, and two of these bonds may be broken for each pair of electrons contributed by a single cage. If cage bonds themselves are broken, the number which may be broken per extra electron pair per cage depends on the number of cages that share a particular linkage: three for the prismatic bonds, b, and two for the pyramidal bonds, c. The structure of β -GaSe (Fig. 3a) shows that the bonds broken are either both the a bonds or all three b bonds, the result being indistinguishable because of the symmetry of the structure. It is important that the number and type of linkages broken compared to the wurtzite structure accords with the number of extra electrons present. Two other differences are of note by comparison with the wurtzite geometry. First, after bond fission, each of the layers moves with respect to the one beneath it. This clearly aids the packing together of the layers. Second, a re-adjustment of the atom positions within the cage has occurred. The Se atoms only occupy the apical positions of the wurtzite cage fragment. GaTe has a related structure with the Te atoms in the apical sites but with some other differences.

A structure where different linkages of the wurtzite cage are broken is that of wolfsbergite $CuSbS_2$. If we write this as $Cu(I)Sb(III)S_2$ to satisfy divalent sulphur then the copper contributes one 4s electron to the electron count. Antimony contributes five electrons and this leads on average to nine electrons per AS unit that is one more than in ZnS and makes this structure iso-electronic with GaSe. Wolfsbergite is a superstructure of wurtzite and we can readily see from Fig. 3b that two pyramidal SbS linkages are broken.

Again the sites of lowest coordination, where the breakage has occurred contain the most electronegative species. This observation accords with a general feature of molecular cage compounds. The sites of lowest coordination are invariably those carrying the highest charge. Just as in discussions⁸ of the forces determining ligand site preferences in molecular compounds the most electronegative ligand will prefer the sites of highest electron density. We do not know at present why wolfsbergite and β GaSe have different structures. In enargite Cu_3AsS_4 ($Cu(I)Cu_2(II)As(III)S_4$) there is half an electron on average per AS formula unit. This is enough to break only one extra cage bond of the wurtzite unit but is insufficient to produce a separation between two halves of the structure as in $CuSbS_2$ or GaSe (we need to break two bonds at least) and the result is a distorted structure based on wurtzite, in which the atoms are displaced from their ideal positions. Instead of regular tetrahedral coordination a range of Cu-S and As-S bond lengths are found. If the structure is regarded as being made up of $Cu_3(I)As(V)S_4$ then the structure is an electron precise one (four electrons on average per AS formula unit) and an undistorted tetrahedral superstructure is expected.

With 10 electrons per formula unit there are now sufficient extra electrons to break two bonds per cage. In addition to the two extra-cage linkages broken in GaSe, the three vertical linkages, b, may also be broken. (Note that each vertical cage linkage is shared by three cages, so that one extra pair of electrons contributed per cage is sufficient for fission of all three.) The As structure is just this and again the layers are moved with respect to each other to aid in packing (Fig. 4a). With 12 electrons one more bond may be broken per cage and the result is the chain structure (Fig. 4b) found for elemental Se or Te. With 14 electrons another bond may be broken and the result is a lattice composed of molecular I_2 units, still retaining (Fig. 5) in fact a geometrical resemblance by packing of the diatomic units, to the parent structures. The almost trivial case of Xe with 16 electrons has no bonds between the atoms and a van der Waals solid results.

Another of the systems that we could have considered (which we will examine in more detail elsewhere) starts with the eight electron rocksalt (NaCl) structure and observes the gradual breaking of bonds as more electrons are added. Thus the layer structures of arsenic and black phosphorus with two extra electrons built from three coordinate atoms are neatly viewed as arising through two different bond-breaking patterns of the

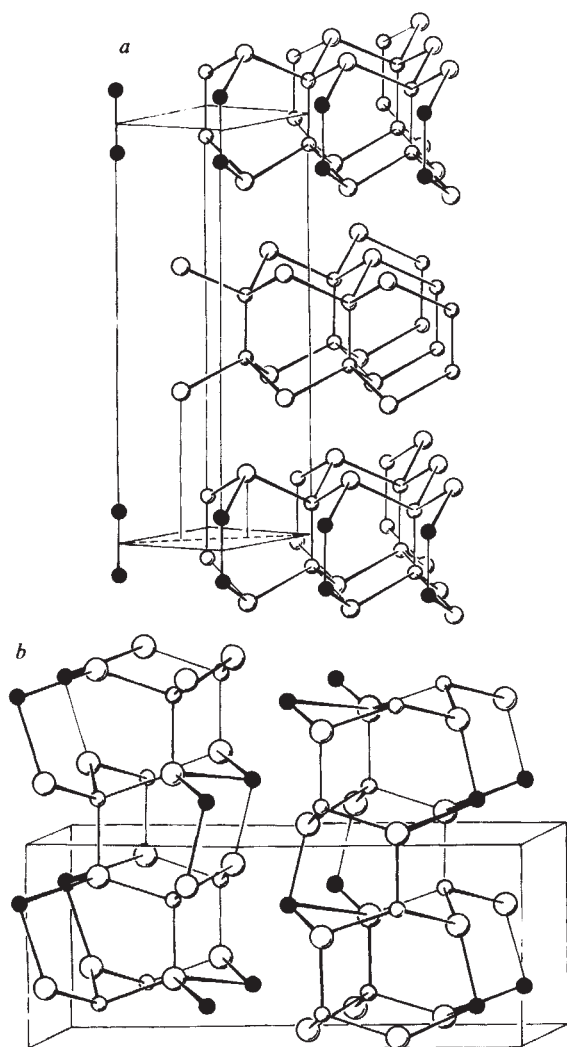


Fig. 3 a, The β GaSe structure; b, the wolfsbergite ($CuSbS_2$) structure. Large circles Se, small circles Ga or Cu, and small full circles, Sb.

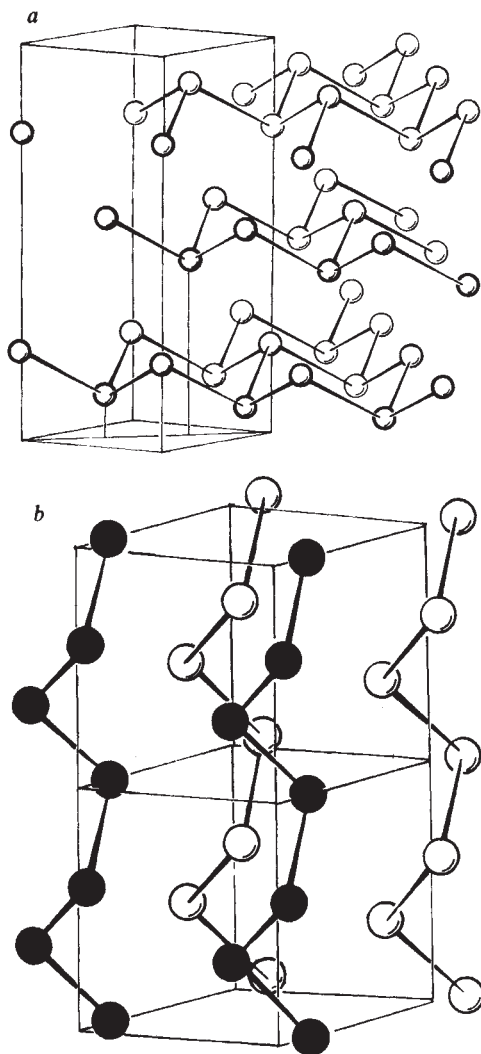


Fig. 4 a, The As layer structure; b, the chain Se, Te structure.

rocksalt structure. TlF and InCl also with two extra electrons have distorted rocksalt-structures where in the latter for example the chlorine atoms are coordinated by three In atoms at 2.91 Å and three at 3.51 Å. InS, with only one extra electron has a structure halfway between that of NaCl and black phosphorus, is a geometry difficult perhaps to understand any other way. Complex sulphur containing minerals with a non-integral number of electrons per formula unit are also approachable using our scheme. Thus marriite PbAgAsS_3 contains on average 9.33 electrons per AS atom pair. The structure of this mineral is a distorted superstructure of rocksalt containing six coordinate lead, three coordinate arsenic and three different sulphur environments. Two are four coordinate and one is five coordinate. On average then per atom pair $1\frac{2}{3}$ 'bonds' of the rocksalt structure are broken; close to what would be predicted by interpolation between the results for eight and 10 electron structures. These ideas could be useful as a way to rationalise complex sulphide and sulphosalts structures in general. A similar analysis applies to silicate structures. SiO_2 itself exists in several modifications, the β -cristobalite and tridymite arrangements being closely related (via an oxygen atom spacer) to those of zincblende and wurtzite. It is also found, with Al^{3+} substituted for Si and a cation present to maintain electrical neutrality, as feldspars and zeolites which are also three dimensional networks. (SiO_2 has 5.33 electrons per atom.) With 5.71 electrons per atom ($(\text{Si}_2\text{O}_5)^{2n-}$ units are found which give rise to double chains (such as gillespite) and sheet structures (such as micas). With 5.87 electrons per atom ($(\text{Si}_4\text{O}_{11})^{6n-}$ is reached, and

double chains (such as amphiboles) are found. With 6.00 electrons, the structure is broken up further ($(\text{SiO}_3)^{2n-}$) and linear or cyclic single chains are found as in benitoite and the pyroxenes respectively. With 6.4 electrons per atom isolated SiO_4^{4-} tetrahedra are found as in the orthosilicates such as the olivines. This neatly shows that the electron rich species are made up of increasingly smaller units.

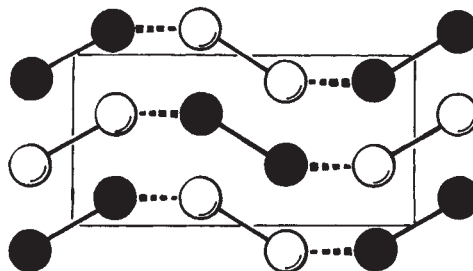


Fig. 5 The I_2 unit cell. The dashed lines show the similarity to the Se, Te spirals of Fig. 4b.

Extra electrons may be introduced in other ways such as inclusion of an alkali or alkaline earth metal into the structure. Thus CaSi_2 (10 electrons per Si_2 unit) has a rumpled sheet structure just like As, and CaSi (12 electrons per Si_2 unit) has zigzag chains just like Se or Te.

Thus, whereas it has been known for a long time that iso-electronic species often have very similar crystal structures we have shown here how the structure adopted may be simply derived. It would be an important step in the classification of the crystalline state if we could relate in general the structure of a particular solid to a reference geometry by consideration of its electronic configuration. We anticipate that this would only work in those systems where covalent or directional forces are important.

Electron counting schemes

A survey of solid state structures based on the zincblende ZnS structure immediately shows some similarities to Wade's molecular scheme. The CdIn_2Se_4 and CdAl_2S_4 species crystallise as defect zincblende structures. (Fig. 6). We may represent them as $\text{Cd}\square\text{In}_2\text{Se}_4$ and $\text{Cd}\square\text{Al}_2\text{S}_4$ where $\square$ = vacancy to emphasise their stoichiometric relationship to ZnS . On average per AS or ASe unit ($\text{A} = \text{Cd}, \text{In}, \text{Al}$ or $\square$) there are eight electrons, just as in ZnS and an average of 12 electrons per 'zincblende' cage. This cage structure is neatly regarded as a *nido* ZnS structure. The zincblende structure of Fig. 2 would, of course, be regarded as a *closo* structure. Examination of a model of this system shows that there are three different zincblende 'cages'. One of these has no atoms missing, the second has two and the third has one.

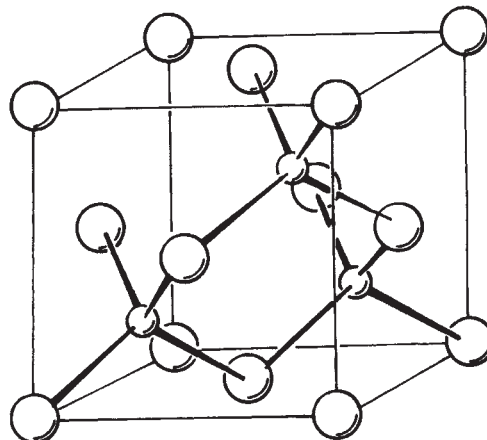


Fig. 6 The defect zincblende structure of CdIn_2Se_4 . Large circles Se.

So the structure contains, *closo*, *nido* and *arachno* cages. There are many other examples.

In fact solid state chemists have used electron counting ideas similar to these for many years, which we may formulate as four principles.

(1) Starting with a close packed cubic array of 'anions', filling of the various holes gives the following structures: zincblende (one tetrahedral hole), CaF_2 (two tetrahedral holes), NaCl (one octahedral hole) and so on.

(2) By substituting different cations with the same or average number of electrons in an ordered fashion a superstructure may be built up.

(3) By substituting cations with more electrons per atom some of the sites of the structure remain vacant in an ordered fashion.

(4) By substituting cations with less electrons per atom filled up derivatives are obtained.

Rule (1) is the solid state analogue of the description of the deltahedra on which the molecular structures are based.

Rule (2) has no molecular analogue in the strictest sense as superstructures are essentially a solid state concept but the best

analogy is that similar structures with the same number of skeletal atoms are obtained on substitution of, for example, a BH group by an $\text{Fe}(\text{CO})_3$ group.

Rule (3) immediately recalls the occurrence of *closo*, *nido* and *arachno* molecular species.

Rule (4) has its molecular counterpart in cluster compounds with a central atom for example, $\text{Ru}_6(\text{CO})_{17}\text{C}$. The analogy between the two schemes applied to molecular species and extended arrays is then very striking.

Received 23 October 1978; accepted 13 March 1979.

1. Mingos, D. M. P. *Nature Phys. Sci.* **236**, 99–102 (1972).
2. Wade, K. *Nature Phys. Sci.* **240**, 71–74 (1972).
3. Wade, K. *Adv. Inorg. Chem. Radiochem.* **18**, 1–66 (1976).
4. Williams, R. E. *Adv. Inorg. Chem. Radiochem.* **18**, 67–142 (1976).
5. Wells, A. F. *Structural Inorganic Chemistry* (4th edn) (Oxford University Press, 1975).
6. Wyckoff, R. W. G. *Crystal Structures* (Wiley, New York 1968).
7. Pearson R. B. *The Crystal Chemistry and Physics of Metals and Alloys* (Wiley, New York, 1972).
8. Rossi, A. R. & Hoffmann, R. *Inorg. Chem.* **14**, 365–374 (1975).
9. Burdett, J. K., Hoffmann, R. & Fay, R. C. *Inorg. Chem.* **17**, 2553–2567 (1978).
10. Parthe, E. *Crystal Chemistry of Tetrahedral Structures* (Gordon and Breach, New York, 1964).

The ovalbumin gene region: common features in the organisation of three genes expressed in chicken oviduct under hormonal control

A. Royal*, A. Garapin*, B. Cami*, F. Perrin†, J. L. Mandel†, M. LeMeur†, F. Brégégère*, F. Gannon†, J. P. LePennec†, P. Chambon† & P. Kourilsky*‡

* Groupe de Biologie Moléculaire du Gène, ERA CNRS 201, Unité de Génie Génétique, Institut Pasteur, Paris, France

† Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS et Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médecine, Strasbourg, France

Two large DNA fragments overlapping the chicken ovalbumin gene have been isolated by molecular cloning. Analysis of these fragments provided a map of a 46,000-base pair region of the chicken genome. This region contains the complete ovalbumin gene (including its mRNA leader-coding sequence) and at least two other genes of unknown function. All three genes are orientated in the same direction and their expression in chicken oviduct is under hormonal control. The three genes share some sequence homologies, suggesting that duplications have occurred in the ovalbumin gene region in the course of evolution.

IN the chicken oviduct, the transcription of several genes, including ovalbumin, is strongly stimulated by steroid hormones^{1–4}. Knowledge of the structure of these genes should help in the understanding of hormonal control and gene regulation in higher organisms. Following the discovery that the ovalbumin gene is split^{5,6} we^{7–11} and others^{12–14} have analysed its organisation using various methods, including molecular cloning of chicken DNA fragments, and demonstrated the presence of seven introns (intervening sequences) in the ovalbumin gene (Fig. 1c; see ref. 15 for review). Evidence for a 47-nucleotide-long leader sequence at the 5' extremity of ovalbumin mRNA (ov-mRNA) has been obtained^{10,14,16}. The leader-coding sequence is not present in any of the gene fragments which we had cloned previously¹⁰.

‡ To whom correspondence should be addressed.

In an attempt to isolate the ov-mRNA leader-coding sequence and to characterise the surroundings of the ovalbumin gene, we decided to clone large DNA fragments overlapping the gene. We report here the isolation of two DNA fragments which cover a 46-kilobase region of the chicken genome. This region contains, in addition to the complete ovalbumin gene including its leader-coding sequence, two other genes which are both expressed in the oviduct under hormonal control, and which share some sequence homologies with the ovalbumin gene.

Isolation of DNA fragments overlapping the ovalbumin gene

We previously^{7,8,11} cloned four *EcoRI* chicken DNA fragments ('a', 'b', 'c' and 'd'), which carry most of the ov-mRNA coding sequences (see Fig. 1c and ref. 10). We next attempted to isolate adjacent regions by cloning, in λ WES (ref. 17), chicken DNA partially digested by *EcoRI* (A.G. *et al.*, unpublished). As this approach failed, we undertook to develop a new cloning system capable of propagating large pieces of DNA (F.B. *et al.*, in preparation). Before these experiments were completed, similar cloning vehicles, named 'cosmids', were described by Collins and Hohn¹⁸. Cosmids are plasmid vehicles which can be packaged *in vitro* into phage λ heads. In this reaction, recombinant molecules are produced in preference to parental molecules and are efficiently converted into infectious particles which can then inject their DNA into bacteria, thus giving rise to bacterial clones harbouring recombinant plasmids. As an efficient screening method was a prerequisite for the isolation of genes from complex genomes by this method, we developed an appropriate *in situ* colony hybridisation technique¹⁹.

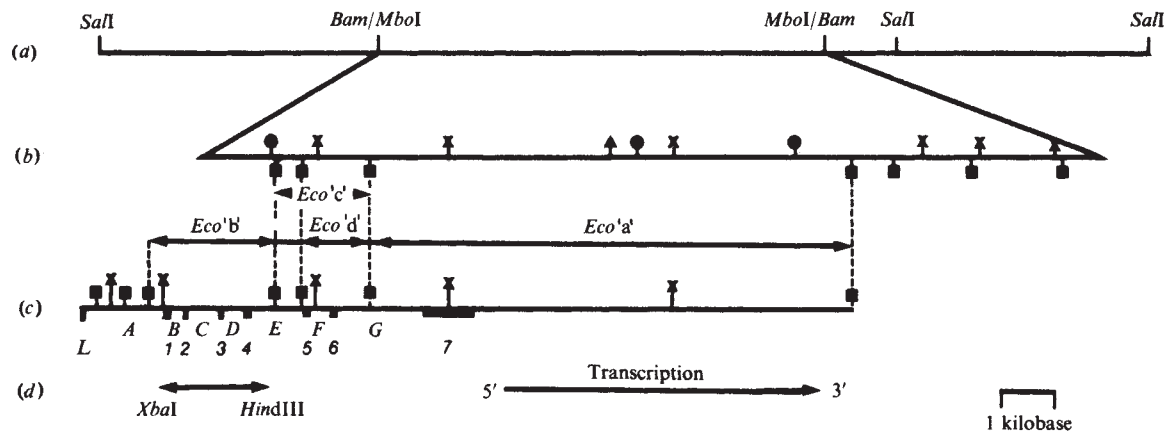


Fig. 1 Restriction enzyme map of the cloned pAR1 DNA. pAR1 was constructed as follows: chicken erythrocyte DNA was partially digested with *Mbo*I and fractionated by sedimentation in a sucrose gradient (5–40% sucrose in 10 mM Tris-HCl pH 8, 1 mM EDTA, 1 M NaCl; sedimentation for 18 h at 25,000 r.p.m. at 4 °C in a SW41 Beckman rotor). Fractions were analysed by agarose gel electrophoresis and those containing fragments of 15 to 25 kilobases were pooled. DNA was concentrated by ethanol precipitation and 5 µg were ligated to an equimolar amount of pJC74Km DNA digested by *Bam*HI as previously described⁷. (We constructed pJC74Km by *in vitro* recombination of *Eco*RI linearised pJC74 (ref. 20) with a 7-kilobase *Eco*RI fragment of pML2 (ref. 21) containing a kanamycin resistance gene.) *In vitro* packaging¹⁸ yielded 250,000 recombinants. After infection of *E. coli* strain 1,400 (ref. 19) bacteria were spread on kanamycin plates and screened by *in situ* colony hybridisation with an ovalbumin-specific ³²P-labelled probe (*Xba*I–*Hind*III probe, shown in line *d*, and derived from cloned ovalbumin *Eco*RI fragment 'b'). A total of 25,000 colonies was examined and one positive clone was obtained. pAR1 plasmid was re-isolated three times. A restriction map of pAR1 was constructed by analysing restriction enzyme digests of pAR1 DNA. *a*, Linear map of pAR1 showing the 17-kilobase *Mbo*I chicken DNA fragment integrated in the *Bam*HI site of pJC74Km. *b*, Enlargement of the 17-kilobase chicken DNA fragment, showing the restriction sites for *Bam*HI (▲), *Eco*RI (■), *Kpn*I (●) and *Xba*I (×). *c*, Map of the ovalbumin gene^{10,16} showing the introns (identified by letters A–G), the exons (numbered 1–7) and the leader-coding sequence (L). The boundaries of *Eco*RI fragments 'a', 'b', 'c' and 'd' are indicated above the line. *d*, Region corresponding to the ovalbumin-specific probe *Xba*I–*Hind*III. Also shown are the direction of transcription and a scale for lines *b* and *c*.

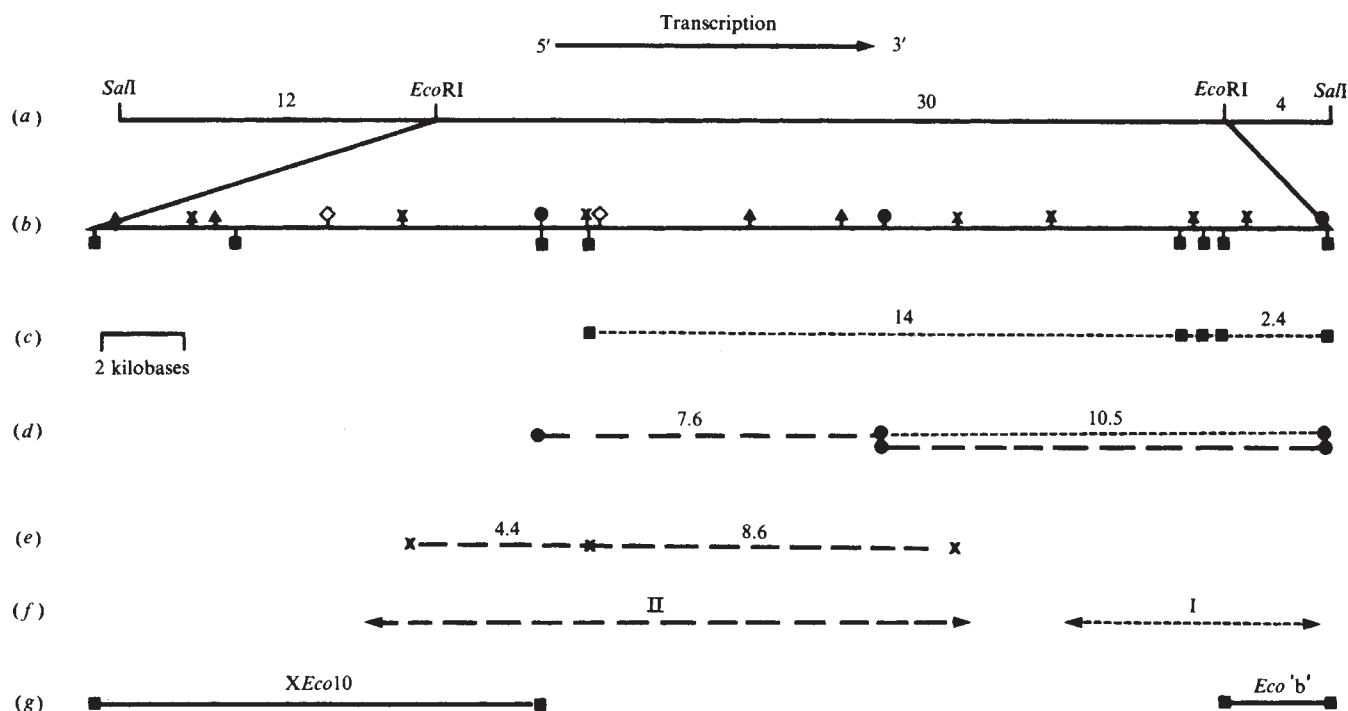


Fig. 2 Restriction enzyme map of the cloned pAR2 DNA. Recombinants were constructed as described in Fig. 1, except that chicken DNA was partially digested with *Eco*RI, and that fragments of more than 25 kilobases (25 µg) were pooled and recombined *in vitro* with 15 µg of *Eco*RI-linearised pJC74 in a final volume of 140 µl. 2 µg of recombinant DNA were packaged *in vitro*. After infection of strain 1,400 and spreading on ampicillin plates, 20,000 colonies were screened with the *Xba*I–*Hind*III probe (see Fig. 1). One positive clone was obtained. Its restriction map was derived from a series of digests performed on whole pAR2 DNA or isolated (subcloned) fragments. It was completed by the identification of fragments which hybridise with an *XEco*10 probe (see below and text) or the ovalbumin probe *Xba*I–*Hind*III (see Fig. 1). *a*, Map of pAR2 DNA linearised with *Sal*I, showing the location of the 30-kilobase insert within the vector sequences. The sizes (in kilobases) of the vector or chicken DNA fragments are indicated above the line. *b*, Map of the *Bam*HI (▲), *Eco*RI (■), *Hpa*II (◇), *Kpn*I (●) and *Xba*I (×) sites in the cloned fragment. *c*, *d* and *e*, maps of the cellular DNA fragments detected in the Southern transfer experiment described in Fig. 3 with either probe I (-----) or probe II (—) (see Fig. 3). Digestion by *Eco*RI (*c*), *Kpn*I (*d*) or *Xba*I (*e*). The size (in kilobases) is indicated above each fragment. *f*, Probes used for comparison of pAR2 to cellular DNA (Fig. 3). Probe I corresponds to a mixture of the 4.5-kilobase *Pst*I and the 3.2-kilobase *Eco*RI fragments in the ovalbumin leader region, isolated from λC4-ov5 (ref. 16). Probe II corresponds to a 14-kilobase long *Hha*I fragment of pAR2 which was subcloned by blunt end ligation into pBR322. *g*, Location of the ovalbumin *Eco*RI fragment 'b' and of the cloned 10.5 kilobase *Eco*RI fragment *XEco*10. This fragment, which was cloned independently in λWES, was selected after hybridisation with the ovalbumin *Hhaov* probe⁸ and will be described elsewhere (J.L.M. et al., in preparation). The scale (for lines *b*–*g*) is given in line *c* in kilobases.

Two cosmids were used: pJC74 (from J. Collins) which can accept foreign DNA fragments of up to 32 kilobases²⁰ and pJ74Km, a kanamycin-resistant derivative which we constructed (see Fig. 1 legend) and which can accept DNA fragments of up to 25 kilobases. Chicken erythrocyte DNA, partially digested by *Mbo*I and sized on a sucrose gradient, was ligated to pJC74Km DNA made linear using *Bam*HI. *Bam*HI and *Mbo*I have identical cohesive ends which allow their ligation. In other experiments, chicken erythrocyte DNA, partially digested by *Eco*RI, was sized on sucrose gradients, and recombined with *Eco*RI-linearised pJC74 DNA. After *in vitro* packaging, infection of *recA* strain 1400 (ref. 19), and screening by *in situ* hybridisation with a ³²P-labelled probe corresponding to part of ovalbumin *Eco*RI fragment 'b' (see Fig. 1d), we isolated two clones, from which the recombinant plasmids pAR1 (derived from pJC74Km) and pAR2 (derived from pJC74) were purified.

The cloned sequences cover 46 kilobases in the region of the ovalbumin gene

Several restriction enzymes were used to construct the maps of pAR1 and pAR2 shown in Figs 1 and 2. Digestion of pAR1 DNA by *Eco*RI yielded eight fragments, including a 9.2-kilobase and a 1.3-kilobase fragment, which co-migrated with the cloned *Eco*RI fragments 'a' and 'd' of the ovalbumin gene^{7,11} (see Fig. 1b, c). The part of the map containing these two fragments confirmed our previous data on the ovalbumin gene organisation^{5,7-11}. There was no fragment of the size of *Eco*RI fragment 'b', but restriction mapping indicated that about half of its sequence was present, as expected as pAR1 was isolated using a probe corresponding to that region. This was further demonstrated by hybridisation of pAR1 DNA with ov-mRNA. Under the electron microscope (not shown), DNA loops corresponding to introns D, E, F and G were observed, in addition to a non-hybridised RNA tail corresponding to exons 1 and 2. From these experiments we conclude that pAR1 contains about 17 kilobases of chicken DNA, starting from the *Mbo*I site located within intron C (refs 8, 10), and ending 12.5 kilobases downstream from the 3' end of the ov-mRNA coding sequence (Fig. 1).

Digestion of pAR2 DNA by *Eco*RI yielded a series of fragments among which the ovalbumin *Eco*RI fragment 'b' (which contains exons 1 to 4) could be identified in a Southern transfer experiment²² using an ovalbumin double-stranded cDNA probe (*Hha*ov, see refs 5, 23), while fragments 'a', 'c' and 'd' (which contain exons 5, 6 and 7) could not be detected (not shown). Unexpectedly, a 3.5-kilobase *Eco*RI fragment and a 2.4-kilobase *Bam*HI fragment, which were not predicted from the ovalbumin gene map, also reacted with the *Hha*ov probe (see below and Fig. 6). Similar *Eco*RI and *Bam*HI fragments which cross-hybridise with the *Hha*ov probe, but do not belong to the ovalbumin gene, have been detected previously in cellular DNA⁹. This suggests that pAR2 contains some of these ovalbumin-related sequences. The restriction map in Fig. 2 (a and b) indicates that pAR2 contains about 30 kilobases of chicken DNA located on the 5' side (with respect to transcription) of the ovalbumin gene (see Figs 1c, 2g and ref. 16). As will be shown below, there are sequence homologies between the ovalbumin gene region which is on the 3' side of fragment *Eco* 'b' and two other genes, X and Y, which are present in pAR2.

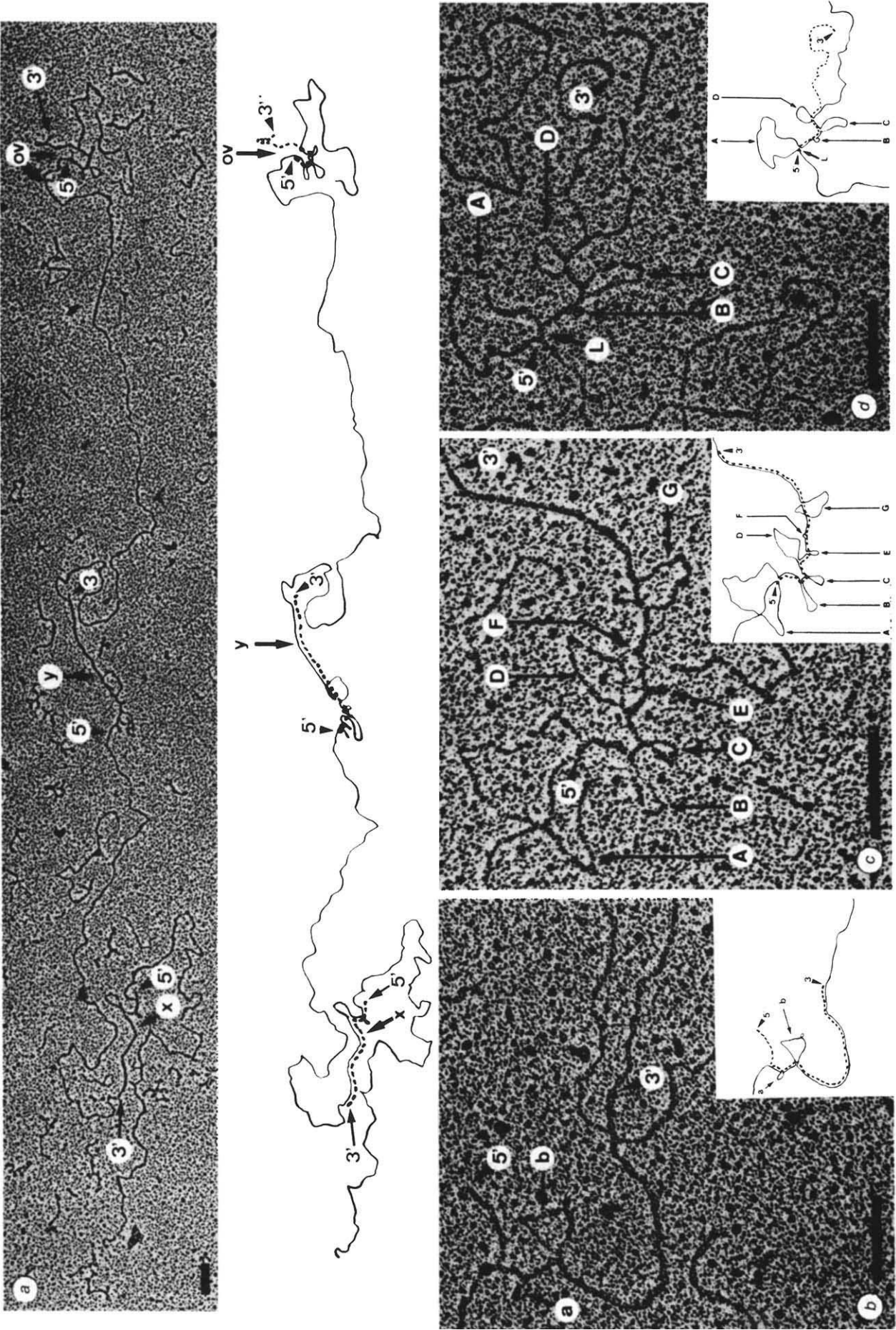
In view of the large size of the DNA fragment which was integrated into pAR2 and because it is the first example of the use of cosmids for cloning eukaryotic DNA, it was important to demonstrate that this cloned fragment exists as such in genomic DNA. We have compared the hybridisation patterns of restriction enzyme digests of pAR2 and chicken oviduct DNA in Southern transfer experiments. Two probes were used for hybridisation (Fig. 2f, g). Probe I covers 6 kilobases in the ovalbumin *Eco*RI fragment 'b' region. Probe II covers 14 kilobases in the middle of the cloned DNA. As shown in Fig. 3, the hybridisation patterns of pAR2 and oviduct DNAs with probe I

are identical in digests made with *Eco*RI (lanes 2 and 6) or *Kpn*I (lanes 4 and 8), or in double digests made with *Eco*RI and either *Bam*HI (lanes 1 and 5) or *Kpn*I (lanes 3 and 7). There are two *Xba*I fragments of similar size in pAR2 and oviduct DNA which hybridise to probe II (lanes 10 and 12). This probe also reacts with three fragments in *Kpn*I digests of both DNAs (lanes 9 and 11). Two of these fragments (bands b and c) have similar sizes in pAR2 and oviduct DNAs. In pAR2, a third fragment of about 26 kilobases which contains 16 kilobases of vector sequences is observed (band a, lane 11), while the corresponding fragment in cellular DNA has a size of about 18 kilobases and contains an



Fig. 3 Restriction enzyme mapping of cloned pAR2 DNA and of the corresponding chromosomal DNA region. pAR2 DNA or oviduct DNA were digested with various restriction nucleases and electrophoresed in 0.7% agarose gels (10 µg of cellular DNA or 10 ng of pAR2 DNA per slot). Transfer on to nitrocellulose filters, hybridisation conditions and autoradiography were as previously described⁵. Digests which were to be analysed with the same probe were all electrophoresed on the same gel, but cellular DNA and pAR2 DNA slots were transferred on to separate filters. Hybridisation of oviduct DNA (lanes 1-4) or pAR2 DNA (lanes 5-8) with nick-translated probe I (see Fig. 2). Fragment sizes are given in kilobases. Lanes 1 and 5: digestion with *Bam*HI and *Eco*RI, a = 7.9, b = 2.4, c = 0.5. Lanes 2 and 6: digestion with *Eco*RI, a = 14, b = 2.4, c = 0.5. (The band above fragment 'b' which appears in lanes 1 and 2 is an artefact.) Lanes 3 and 7: digestion with *Eco*RI and *Kpn*I, a = 7.2, b = 2.2, c = 0.5. Lanes 4 and 8: digestion with *Kpn*I, a = 10.5. Hybridisation of oviduct DNA (lanes 9 and 10) or pAR2 DNA (lanes 11 and 12) with nick-translated probe II (see Fig. 2). Lanes 9 and 11: digest with *Kpn*I, a (lane 9) = 18, a (lane 11) = 26, b = 10.5, c = 7.6. Lanes 10 and 12: digest with *Xba*I, a = 8.6, b = 5.1, c = 4.4. Band b (lane 12) corresponds to the *Xba*I fragment adjacent to fragment 'c' (at its left, see Fig. 2b, e). As fragment 'b' hybridises only to a very small portion of probe II, it was not detected in the cellular DNA.

additional 8-kilobase segment not present in pAR2 (band a, lane 9). As shown in Fig. 2c, d and e, the fragments hybridising to probes I and II and which are common to oviduct and pAR2 DNA, cover a 24-kilobase segment. In addition, the 10.5-kilobase region of the cloned DNA which is distal to the ovalbumin sequence, and which overlaps this 24-kilobase segment, has been found to be identical to an independently cloned chicken DNA fragment XEco10 (see Fig. 2g) which has been studied separately (J.L.M. *et al.*, in preparation). From all these results we conclude that the entire 30-kilobase region cloned in pAR2 corresponds to a single continuous segment of cellular DNA, which has not suffered any major rearrangement during cloning.



pAR2 contains the ovalbumin leader-coding sequence

As pAR2 contains a very large segment of DNA on the 5' side of the ovalbumin gene, we expected that it should have the ov-mRNA leader-coding sequence¹⁰. We have analysed by electron microscopy the hybrid molecules formed between ovalbumin mRNA and heat-denatured pAR2 DNA, previously linearised by *SalI* (Fig. 4d). ov-mRNA hybridises over 620 ± 67 nucleotides in five regions separated by four single-stranded DNA loops. Three of these loops have sizes and locations corresponding to introns B, C and D of the ovalbumin gene which allows the orientation of the molecule with respect to transcription^{8,9}. The fourth DNA loop which is at the 5' end of ov-mRNA is $1,606 \pm 150$ nucleotides long (Fig. 4d). This loop corresponds to intron A, which was expected to separate exon 1 from the sequence coding for the mRNA leader¹⁰. Indeed both the length of loop A and the location of several restriction sites in the corresponding region of pAR2 agree with data obtained on two other independently isolated clones where the presence of the leader-coding sequence at the 5' end of the loop was unequivocally demonstrated by DNA sequencing¹⁶. The free mRNA tail, which is about 1,200 nucleotides long, corresponds to exons 5, 6 and 7 which are not present in pAR2.

pAR2 contains two other genes expressed in oviduct of laying hen

In the 46-kilobase DNA region cloned in pAR1 and pAR2, the ovalbumin gene occupies only about 7.7 kilobases¹⁶. This could leave ample space for other genes. As we were interested to know whether genes which share a similar hormonal regulation of expression in chicken oviduct could be clustered, we have hybridised total poly(A) RNA from laying hen oviducts with pAR1 or pAR2 DNAs. No RNA other than ov-mRNA was found to hybridise to pAR1 DNA. In contrast, electron microscopy of the hybrid molecule revealed that, besides ov-mRNA, two other RNA molecules hybridised to the same strand of pAR2 DNA (Fig. 4a), indicating the presence of three genes transcribed in the same direction (Fig. 4a). The polarity with respect to transcription is given by the ov-mRNA-DNA hybrid region which is easily identified by the characteristic array of DNA loops corresponding to introns B, C and D (Fig. 4a, d, see below for the absence of the loop corresponding to intron A in Fig. 4a). We termed these two other genes, X and Y, giving the gene order 5'-X-Y-ovalbumin-3'.

Y RNA is $2,020 \pm 175$ nucleotides long, as measured from the length of the DNA-RNA hybrid regions located in the middle of the cloned segment. The DNA loops indicate the presence of seven introns (A to G) in this gene, separating eight exons (1 to 8) (Figs 4c and 5a, b). Note that with Y RNA, as with ov-mRNA, the DNA loop closest to the 5' end of the RNA is not seen on all hybrid molecules (and is in fact absent for both genes in Fig. 4a). In the case of ov-mRNA this is due to the short size (47 nucleotides) of the leader sequence which does not form a very stable hybrid in the conditions used to form preferentially

DNA-RNA hybrids¹⁶. A similar situation is probably encountered with Y RNA, as its exon 1 is not actually resolved by electron microscopy and its existence is inferred from the presence of loop A (Fig. 4c). The total length of the Y gene (exons + introns) is about 6.5 kilobases, and its 3' end lies about 11.5 kilobases upstream from the ov-mRNA leader-coding sequence.

X RNA is about 2,400 nucleotides long. It hybridises to pAR2 DNA over $1,945 \pm 200$ nucleotides in three exons separated by two DNA loops (introns) (Figs 4b and 5a, b). The presence of a free RNA tail of about 400 nucleotides indicates that the region coding for the 5' end of this RNA is not present in pAR2 DNA. There are about 5.5 kilobases between the 3' end of the X gene and the 5' end of the Y gene. Using these data, genes X and Y were positioned on the restriction enzyme map as shown in Figs 5c and 6B(1).

X, Y and ovalbumin genes share sequence homologies

In the original mapping of pAR2 DNA (see above) we found that, in addition to the ovalbumin gene sequences, some other parts of the DNA hybridised with the ovalbumin-specific *Hha*ov probe^{5,23}. The finding of genes X and Y prompted us to investigate whether the cross-hybridising regions could correspond to these genes.

pAR2 DNA was digested by *Hpa*II and either *Eco*RI or *Bam*HI. *Hpa*II cuts the cosmid DNA into very small fragments, facilitating further analysis of the cloned chicken DNA, which is cut only two times. Digestion of the DNA from pAR2 with *Hpa*II and *Eco*RI yields one band (band c, lane 1, Fig. 6A) which hybridises strongly with the *Hha*ov probe. Two other bands of weaker intensity (bands a and b, lane 1) were also revealed. From the restriction enzyme mapping of pAR2 we can unambiguously identify these bands and show their location in Fig. 6B, line 1. Similarly, in a *Hpa*II and *Bam*HI double digest, hybridisation with the *Hha*ov probe yielded one intense band (band a, lane 2, Fig. 6A), which includes the ovalbumin coding region and two other bands (bands b and c, lane 2), which are located in the regions of the DNA coding for genes Y and X, respectively (see Fig. 6B, line 2). These homologies were further analysed using a probe from the region of gene X (probe V, Fig. 6B, line 5). On digestion with *Hpa*II and *Eco*RI, two bands of hybridisation were obtained (lane 3, Fig. 6A): band a which includes part of gene Y, and band b which contains the region from which the probe was prepared (see Fig. 6B, line 3). Using the same probe with DNA digested by *Hpa*II and *Bam*HI, three bands of hybridisation were found (Fig. 6A, lane 4). Two of these bands (c and d) originate from gene X, whereas band b corresponds to a part of gene Y (see Fig. 6B, line 4). These experiments show that there are homologies between regions of genes X and Y, and that both genes contain sequences which are homologous to some ovalbumin exonic sequences. Even so, these three genes and their RNA transcripts are clearly different, as indicated by the electron microscopic results and by

Fig. 4 Electron micrographs of RNA-DNA hybrid molecules between laying hen oviduct RNA purified on oligo(dT)-cellulose and pAR2 DNA. RNA-DNA hybrids were formed by incubating heat-denatured linear pAR2 DNA ($1.5 \mu\text{g ml}^{-1}$) with oligo(dT)-cellulose-purified RNA ($100 \mu\text{g ml}^{-1}$) at 54.3°C and prepared for electron microscopy⁷. pAR2 DNA was linearised by digestion with *SalI*. In the line drawings, the RNAs are represented as a dashed line and the DNAs as a solid line. Scale bars represent $0.1 \mu\text{m}$. **a**, Electron micrograph of a complete single-stranded pAR2 DNA molecule hybridised with three RNA molecules. Arrows X, Y and ov point to the respective genes. Arrow heads 5' and 3' indicate the 5' and the 3' extremities of the RNAs. The total length of pAR2 DNA is $45,940 \pm 6,600$ base pairs (from the measurements of 20 molecules). The X RNA hybridises to a DNA sequence which lies $10,460 \pm 1,150$ base pairs from the nearest pAR2 DNA end. There are $5,560 \pm 600$ base pairs between the 3' end of the X gene and the 5' end of the Y gene and $11,580 \pm 1,350$ base pairs between the 3' end of the Y gene and the 5' end of the ovalbumin gene. There are $5,330 \pm 540$ base pairs between the DNA sequence which hybridises to ov-mRNA and the end of the molecule. Loops A of the ovalbumin gene (see Fig. 4b) and of the Y gene (see Fig. 4c) are not visible in this picture, presumably because the hybrids responsible for their presence are not very stable (see text). **b**, Electron micrograph of an X RNA-DNA hybrid molecule. The hybrid region is $1,945 \pm 243$ base pairs long (from the measurements of 20 molecules). The free RNA tail is about 450 nucleotides long. The two single-stranded DNA loops 'a' and 'b' are 253 ± 47 and 812 ± 120 nucleotides long, respectively. The hybrid segments are (from 5' to 3') 175 ± 39 , 181 ± 35 and $1,613 \pm 219$ base pairs, respectively. Other symbols as under (a). **c**, Electron micrograph of a Y RNA-DNA hybrid (from the measurements of 20 molecules). The hybrid region is $2,020 \pm 175$ base pairs. The length of the single-stranded DNA loops A to G are $1,729 \pm 180$; 492 ± 130 ; 303 ± 88 ; 812 ± 135 ; 206 ± 86 ; 55 ± 20 and 890 ± 134 nucleotides long, respectively. The hybrid segments (from 5' to 3') are 218 ± 71 ; 93 ± 30 ; 165 ± 40 ; 153 ± 33 ; 167 ± 33 ; 183 ± 41 and $1,320 \pm 148$ base pairs, respectively. Other symbols as under (a). **d**, Electron micrograph of an ovalbumin mRNA-DNA hybrid (from the measurements of 22 molecules). The hybrid region is 620 ± 67 base pairs long. The free RNA tail is about 1,100 nucleotides long. Loops A to D are $1,606 \pm 150$; 204 ± 47 ; 578 ± 95 and 386 ± 85 nucleotides long. The hybrid segments are (from 5' to 3') 213 ± 45 , 76 ± 23 , 171 ± 33 and 162 ± 46 base pairs long. The L arrow points to the leader-coding segment which is revealed by the presence of loop A. Other symbols as under (a).

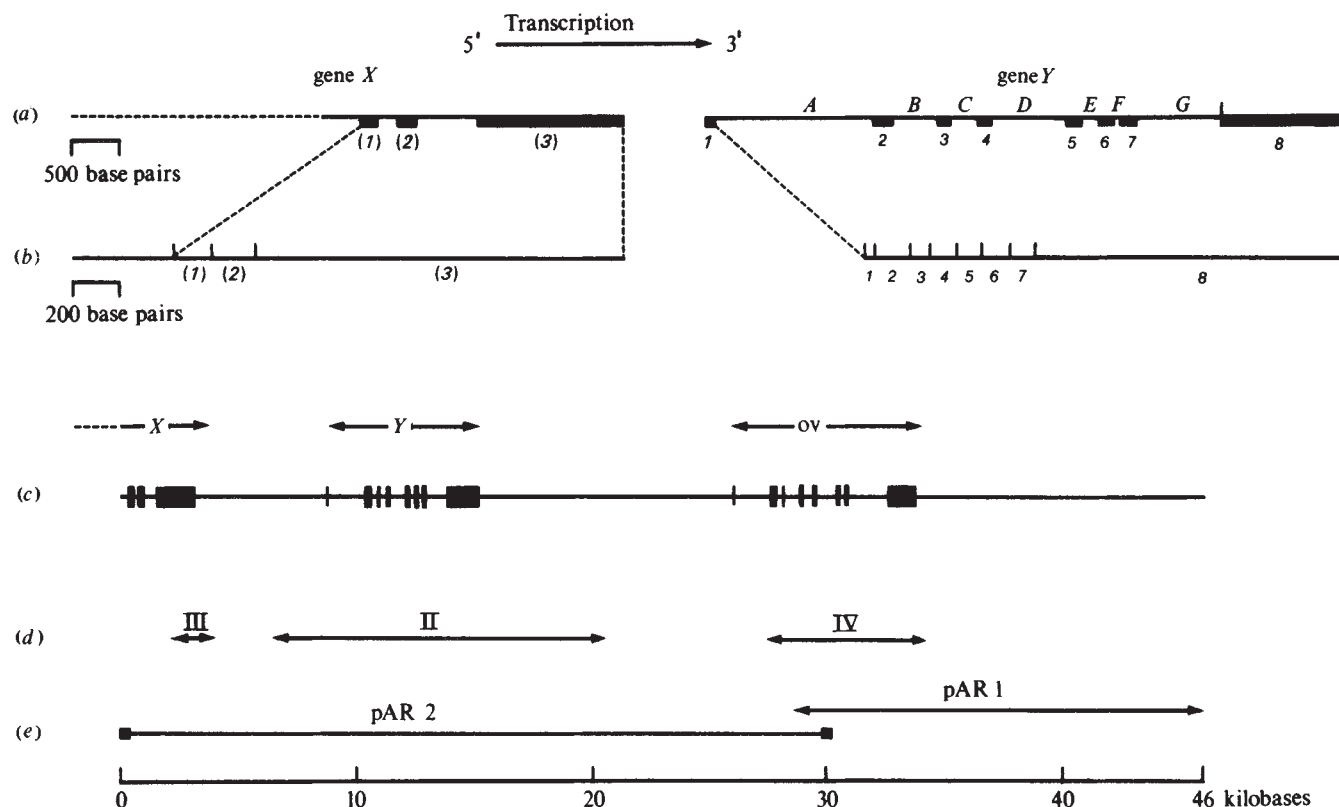


Fig. 5 Map of the ovalbumin gene region. *a*, *b*, Maps of genes *X* and *Y* indicating the respective locations of introns (designated by letters), exons (heavy lines and numbers) which are also shown on the maps of the corresponding RNAs (line *b*) where they are designated by numbers (for gene *X* the numbers are in brackets as definitive numbering of exons requires knowledge of the complete structure of the gene). *c*, Map of the 46-kilobase DNA region contained in pAR1 and pAR2, showing the location of the exons (heavy lines) of *X*, *Y* and ovalbumin (*ov*) genes. *d*, Probes used for the experiment in Fig. 7. Probe II is described in Fig. 2 legend. Probe III is a 1.8-kilobase *Hind*III fragment isolated from *X*Eco10 (see Fig. 2 legend). Probe IV corresponds to the ovalbumin *Eco*RI fragment 'b' and 'c' and to a part of *Eco*RI fragment 'a' (see Fig. 1c). *e*, Limits of the fragments integrated in pAR1 and pAR2 and scale (in kilobases) for lines *c*, *d* and *e*.

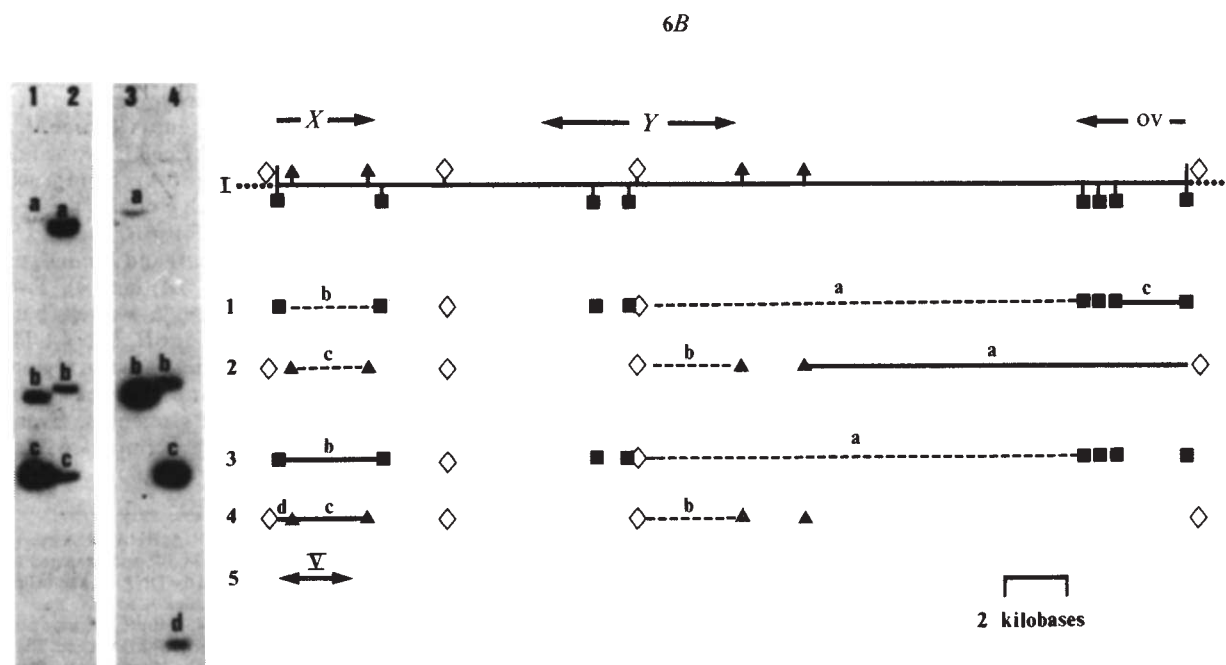


Fig. 6 Sequence homologies between *X*, *Y* and ovalbumin genes. *A*, Digests of pAR2 DNA with *Eco*RI and *Hpa*II (lanes 1 and 3) or *Bam*HI and *Hpa*II (lanes 2 and 4) were electrophoresed on a 0.7% agarose gel, transferred on to nitrocellulose filters and hybridised to probes corresponding to the ovalbumin gene (*Hha*ov: lanes 1 and 2) or to gene *X* (probe V, lanes 3 and 4). The latter probe, shown in line 5 of Fig. 6B is a 2.3-kilobase *Hind*III fragment subcloned from *X*Eco10 (see Fig. 2 legend). The size of the fragments are (in kilobases): lanes 1 and 3, *a* = 14, *b* = 3.4, *c* = 2.4; lanes 2 and 4, *a* = 12.5, *b* = 3.6, *c* = 2.4 and *d* = 0.9. *B*, The location of the fragments hybridising in lanes 1 to 4 of Fig. 6A are indicated in lines 1 to 4, respectively. For a given probe the continuous line shows the fragment(s) which contain(s) part or all of the sequences of the probe, whereas dashed lines indicate additional fragments which cross-hybridise with the probe. The first line (I) gives the location of *Bam*HI (▲), *Eco*RI (■) or *Hpa*II (◇) restriction sites and indicate the regions corresponding to genes *X*, *Y* and ovalbumin. The dotted line corresponds to vector sequences. Line 5: location of probe V and scale.

the lack of hybridisation of the *Y* gene with the *XbaI-HindIII* probe (Fig. 1d) prepared from the ovalbumin *EcoRI* fragment 'b' (not shown).

Expression of *X* and *Y* is under hormonal control

As transcription of the ovalbumin gene is under oestrogen control, we wondered whether the expression of *X* and *Y* is regulated in a similar way. Oviduct RNA from laying hen or oestrogen-withdrawn chicken was fractionated on methyl mercuric hydroxide agarose gels, transferred and immobilised on to diazobenzyloxymethyl paper²⁴. The filters were hybridised with probes specific for each of the three genes.

As shown in Fig. 7, probe IV (see Fig. 5d), specific for the ovalbumin gene, hybridised with an RNA molecule about 1,900 nucleotides long, which is present in laying hen oviduct (lane 3) but cannot be detected in 'hormone-withdrawn' oviduct (lane 1), as expected for ovalbumin mRNA. Probe III (see Fig. 5d) is specific for gene *X* as it does not contain the regions which cross-hybridise with either *Y* or ovalbumin genes (J.L.M. *et al.*, in preparation). An RNA species present in laying hen (Fig. 7, lane 7), but not in 'hormone-withdrawn' oviduct (lane 6) hybridises to this probe. As expected from the above electron microscopic results, the size of this RNA is about 2,300 nucleotides. Gene *Y* is not expressed in the 'hormone-withdrawn' oviduct (Fig. 7, lane 4) as no RNA species hybridised with probe II which contains gene *Y* (see Fig. 5d). In laying hen oviduct (lane 5) we found an RNA species, about 2,000 nucleotides long, which hybridised with probe II. But we cannot state with certainty that it is RNA *Y*, because probe II also cross-reacts with ov-mRNA, which is of similar size. However, the electron microscopic results indicate unambiguously that RNA *Y* is present in the oviduct of laying hens. That the transcription of *X* and *Y* genes is oestrogen-dependent is indicated by the presence of *X* and *Y* RNAs in oviducts following administration of oestradiol to 'withdrawn' chicken (not shown). However, the results shown in Fig. 7 and the frequency of RNA-DNA hybrids observed for *X* and *Y* genes by electron microscopy suggest that the concentration of *X* and *Y* RNAs is much lower than that of ov-mRNA.

Organisation of the chicken genome in the vicinity of the ovalbumin gene

Cloning of eukaryotic DNA fragments using cosmid vectors has not been described previously. Our experiments show that this system, in conjunction with an appropriate *in situ* hybridisation method¹⁹, can be used successfully. Available phage vectors cannot propagate more than about 20 kilobases^{17,25}. Using pJC74 cosmid, we have cloned a 30-kilobase DNA segment of the chicken genome. Shorter cosmid vectors should permit the cloning of DNA fragments of up to 40–45 kilobases. As such large DNA fragments have a greater chance of containing repeated sequences and might be unstable during cloning, we used a *recA*⁻ bacterial host (some phage vectors, including λ Charon 4A (ref. 25), would not grow on such hosts). The efficiency of the cosmid vector system is such^{18,20} that it is possible to construct a genomic library from a few microgrammes of eukaryotic DNA. Enrichment for a particular gene sequence before cloning seems to be unnecessary and even undesirable, as it might lead one to discard fragments of unexpected interest, such as genes of related sequences.

In one clone we found two other genes (*X* and *Y*) closely linked to the ovalbumin gene. As the existence of genes *X* and *Y* was revealed by electron microscopy of hybrid molecules between total oviduct poly(A)⁺ RNA and the cloned DNA, we would not have detected genes coding for minor species of oviduct RNAs or RNAs expressed in other tissues. Assuming that no other genes are present in the cloned DNA, the length of the intergene regions (5–10 kilobases) and of the genes themselves (6–8 kilobases), is such that there would be one gene every 10 to 15 kilobases of DNA.

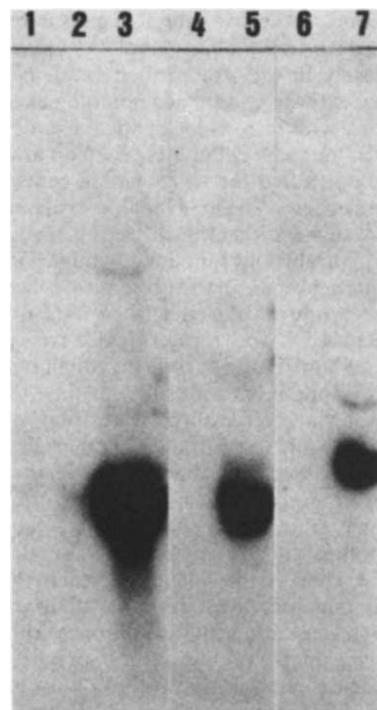


Fig. 7 Detection of *X* and *Y* RNAs in oviducts of laying hen and 'withdrawn' chicken. Female chicks aged 5 days were injected daily with 1 mg of oestradiol benzoate for a period of 10 days (primary stimulation). Hormone treatment was then stopped and 4 weeks later the magnum portion of the oviduct was removed ('withdrawn' oviduct). Total RNA from laying hen or hormone-withdrawn oviduct was prepared according to a method developed by Dr F. Rougeon (personal communication). RNA samples (20 μ g per slot) were separated by electrophoresis on 1% agarose gels containing 10 mM methyl mercury hydroxide. The RNA was then transferred to diazobenzyloxymethyl paper which was hybridised to ³²P-labelled probes and washed²⁴. Autoradiography was with a Kodirex film (Eastman Kodak) for 4 days. Lanes 1–3: hybridisation to 2.5 $\times 10^6$ d.p.m. of ³²P-labelled probe IV (see Fig. 5d); lanes 4 and 5: hybridisation to 2.5 $\times 10^6$ d.p.m. of ³²P-labelled probe II (see Figs 2f and 5d); lanes 6 and 7: hybridisation to 4 $\times 10^6$ d.p.m. of ³²P-labelled probe III (see Fig. 5d); lanes 1, 2, 4 and 6: 20 μ g of total oviduct RNA from 'withdrawn' chicken; in lane 2, 0.1 μ g of laying hen oviduct RNA was added; lanes 3, 5 and 7: 20 μ g of total RNA from laying hen oviducts.

As *X*, *Y* and ov-mRNAs are produced in differentiated cells, could reorganisation of the genome in this region occur during differentiation, as in the case of immunoglobulin genes? (See refs 26 and 27 and refs therein.) Our comparison of the restriction map of the erythrocyte DNA cloned in pAR2 and of either oviduct DNA (Fig. 3) or erythrocyte DNA (results not shown) indicates that there is no major rearrangement.

There are striking resemblances (Figs 4 and 5c) in the organisation of these neighbouring genes, which are transcribed from the same strand and are under oestrogen control. More than half of the RNA-coding sequences are found uninterrupted at the 3' end of all three genes. For both the ovalbumin and *Y* genes there are seven introns and the eight regions coding for the mature RNAs are very similar in size. In addition, the first intron (intron A) is about 1,600 nucleotides long in the two genes and separates a very short RNA-coding sequence from the neighbouring exon. It will be interesting to find whether exon 1 of gene *Y* is equivalent to the mRNA leader-coding sequence of the ovalbumin gene¹⁶. The above similarities and the fact that genes *X*, *Y* and ovalbumin share partial sequence homologies strongly suggest that they could have arisen by at least partial duplications from a common precursor. In our previous work (refs 7, 9; see also Fig. 1 in ref. 5) we detected a limited number of cellular DNA fragments in the chicken genome, which cross-hybridised to the *Hha*ov ovalbumin cDNA probe, but did not belong to the ovalbumin gene. Genes *X* and *Y* account for the existence of most, if not all, of these fragments (our unpublished results). The presence in the chicken genome of other genes with the same degree of homology with the ovalbumin gene is therefore unlikely.

The organisation of the ovalbumin region is reminiscent of that found for the β -like globin genes^{28,29}. Genes for β - and δ -globin are closely linked, transcribed from the same DNA strand, and show extensive sequence homologies in the mRNA coding sequences, with introns in identical locations. They are both expressed in the same cellular type, albeit at very different rates. It is also known that the two γ -globin genes are found in the same genomic region²⁸. The available evidence for several gene families^{30,31} suggests that the clustering of genes which are related both structurally and functionally might not be uncommon in eukaryotic genomes. In the present case, however, we do not know if the products of genes X, Y and ovalbumin are functionally related, as the former two have not yet been identified (genes X and Y do not code for conalbumin, lysozyme or ovomucoid; our unpublished results).

Could the clustering of the three genes play a part in their expression? They could participate together in a chromatin domain (see ref. 32 and refs therein) or be transcribed and controlled by oestrogens as part of one unit. In prokaryotes, it has been demonstrated that genes of related function can be clustered in operons and transcribed by a single RNA-polymerase molecule. There is no evidence at present that this occurs in eukaryotes. In our case, transcription of all three genes by the same RNA-polymerase molecule would have to be reconciled with the following facts: (1) the largest precursor RNA for ovalbumin so far detected is about 8,000 nucleotides^{33,34}; (2) there are large differences in the levels of the three RNAs (ov-mRNA being the most abundant); and (3) the clustering of the three genes may have no other significance than reflecting the evolutionary process which gave rise to them.

Biohazards associated with the experiments described in this publication were examined previously by the French National Control Committee. The experiments were carried out in L3-B1 conditions in the nomenclature adopted by the French Committee (*Le Progrès Scientifique* No. 191, Nov/Dec 1977). We thank Dr J. Collins for providing pJC74, Dr B. Hohn for providing strains, information and advice for *in vitro* packaging, Dr P. Courvalin for pML2 DNA, and Dr F. Rougeon for T₄ DNA ligase and useful advice. The technical help of Mrs O. Lebaill, A. Hémar, Mrs C. Kutschis, E. Sittler, M. C. Chanal, C. Wasyluk, Mr Garnier and B. Boulay is acknowledged. This work was supported by grants to P.K. from the CNRS (ERA 201,

ATP No. 3558), INSERM (CRAT 76.4.319 and 76.4.465), the DGRST and the Fondation de la Recherche Médicale Française, and by grants to P.C. from the CNRS (ATP 2117), INSERM (CRAT 76.5.462 and 76.5.468) and the Fondation pour la Recherche Médicale Française. J.P. LeP. was supported by a fellowship from the Université Louis Pasteur and A.R. is a post-doctoral fellow of the NCI of Canada. Electron microscope facilities were made available to F.P. by INSERM and CNRS.

Received 29 January; accepted 26 March 1979.

- Palmiter, R. D. *Cell* **4**, 189–197 (1975).
- McKnight, G. S. *Cell* **14**, 403–413 (1978).
- Hynes, N. E., Groner, B., Sippel, A. E., Nguyen-Huu, M. C. & Schütz, G. *Cell* **11**, 923–932 (1977).
- Bellard, M., Gannon, F. & Chambon, P. *Cold Spring Harb. Symp. quant. Biol.* **42**, 779–791 (1977).
- Breathnach, R., Mandel, J. L. & Chambon, P. *Nature* **270**, 314–319 (1977).
- Doel, M. T., Houghton, M., Cook, E. A. & Carey, N. H. *Nucleic Acids Res.* **4**, 3701–3713 (1977).
- Garapin, A. C. *et al.* *Nature* **273**, 349–354 (1978).
- Garapin, A. C. *et al.* *Cell* **14**, 629–639 (1978).
- Mandel, J. L. *et al.* *Cell* **14**, 641–653 (1978).
- Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853–4857 (1978).
- LePenec, J. P. *et al.* *Nucleic Acids Res.* **5**, 4547–4562 (1978).
- Weinstock, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1299–1303 (1978).
- Dugaiczky, A. *et al.* *Nature* **274**, 328–333 (1978).
- Catterall, J. F. *et al.* *Nature* **275**, 510–513 (1978).
- Kourilsky, P. & Chambon, P. *Trends biochem. Sci.* **3**, 244–247 (1978).
- Gannon, F. *et al.* *Nature* **278**, 428–434 (1979).
- Enquist, L., Tiemeier, D., Leder, P., Weisbert, R. & Sternberg, N. *Nature* **259**, 596–598 (1976).
- Collins, J. & Hohn, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4242–4246 (1978).
- Cami, B. & Kourilsky, P. *Nucleic Acids Res.* **5**, 2381–2390 (1978).
- Collins, J. & Bruning, H. J. *Gene* **4**, 85–107 (1978).
- Hershfield, V., Boyer, H. W., Yanofsky, C., Lovett, M. A. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3455–3459 (1974).
- Southern, E. M. *J. molec. Biol.* **98**, 503–518 (1975).
- Humphries, P. *et al.* *Nucleic Acids Res.* **4**, 2389–2406 (1977).
- Alwine, J. C., Kemp, D. J. & Stark, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5330–5334 (1977).
- Blattner, F. R. *et al.* *Science* **196**, 161–169 (1977).
- Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15**, 1133–1145 (1978).
- Seidman, J. G. & Leder, P. *Nature* **276**, 790–795 (1978).
- Flavell, R. A., Kooter, J. M., DeBoer, E., Little, P. F. R. & Williamson, R. *Cell* **15**, 25–41 (1978).
- Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157–1174 (1978).
- McKeown, M. *et al.* *Cell* **15**, 789–800 (1978).
- Klein, J. A. *Rev. Genet.* **8**, 63–77 (1974).
- Chambon, P. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1209–1234 (1977).
- Roop, D. R., Nordstrom, J. L., Tsai, S. Y., Tsai, M. J. & O'Malley, B. W. *Cell* **15**, 671–685 (1978).
- Chambon, P. *et al.* in *From Gene to Protein: Information Transfer in Normal and Abnormal Cells* (Academic, New York, in the press).

Histone genes are clustered with a 15-kilobase repeat in the chicken genome

R. J. Crawford, P. Krieg, R. P. Harvey, D. A. Hewish & J. R. E. Wells

Department of Biochemistry, University of Adelaide, Adelaide, South Australia 5001

Histone mRNA isolated from 5-day-old chick embryos has been used as a template for complementary DNA (cDNA) synthesis. The resultant cDNA, after removal of sequences complementary to rRNA, was used to detect histone genes in adult chicken genomic DNA. Hybridisation data indicate that the histone genes are repeated about 10-fold in the chicken genome. Restriction endonuclease analysis reveals some sequence heterogeneity in these genes. However, the results show that chicken histone genes are clustered with a basic repeat unit of 15 kilobases.

THE organisation of histone genes in sea urchin species has been elegantly demonstrated by mapping restriction endonuclease fragments^{1–3} and by electron microscopy⁴. A prerequisite for these studies was the availability of a probe of sufficient radio-

chemical purity to allow unequivocal detection of histone gene sequences. In the sea urchin system, it was possible to take advantage of the fact that early cleavage embryos synthesise predominantly histone mRNA. Pulse-labelled RNA isolated from these cells, although contaminated with other RNA, was perfectly usable because almost all the radioactivity was contained in histone mRNA.

Histone gene sequences from the sea urchin have also been useful in studying *Drosophila* histone genes. Cross reactivity between sea urchin and *Drosophila* histone coding sequences has led to the isolation and subsequent analysis of cloned *Drosophila* histone genes⁵. Other systems have not been so amenable. Despite intensive efforts, isolation of histone mRNA of comparable radiochemical purity from pulse-labelled *Drosophila* tissue culture cells⁶ or HeLa cells⁷ has been very difficult. Purification of human^{8,9} or *Xenopus*¹⁰ histone mRNA sequences by physicochemical methods has similarly proved

elusive. By comparison, 5-day chick embryos seem to be an excellent source of histone mRNA. When cDNA is prepared from this RNA template and sequences complementary to rRNA are removed, the resulting probe can be used to detect histone genes in the genome.

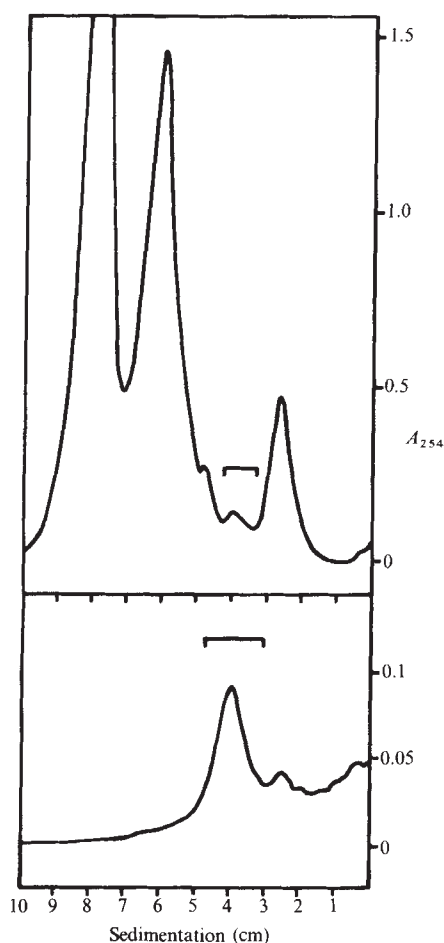


Fig. 1 Isolation of chick histone mRNA. Five-day-old chick embryos were snap-frozen in liquid N_2 , then homogenised in 7 M guanidinium-Cl in 20 mM Tris-Cl, pH 7.5, 1 mM EDTA and 1% (w/v) Sarkosyl in a Dounce homogeniser in a final volume of 30 ml. Total RNA from this homogenate was recovered as material centrifuged through 5.7 M CsCl as described elsewhere¹¹. The clear RNA pellets were gently homogenised in 10 mM Tris-Cl, pH 7.5, 1 mM EDTA, 5% Sarkosyl and 5% phenol, then made 0.1 M in NaCl and phenol extracted with an equal volume of phenol: chloroform (1:1 v/v). RNA from the aqueous phase was collected after precipitation with ethanol, resuspended in 10 mM Tris-Cl, pH 7.4, 1 mM EDTA and 0.5% SDS, heated at 65 °C for 5 min, chilled and centrifuged on 10–40% sucrose gradients at 220,000g for 16 h at 4 °C. The 7–11S RNA was collected and refractionated on a second gradient as shown. Thirty embryos yielded 10–30 μ g of 7–11S RNA.

Characterisation of RNA containing histone mRNA

Total RNA was prepared from 5-day-old chick embryos¹¹, and a discrete 7–11S RNA species was resolved from other major RNAs by sucrose gradient centrifugation (Fig. 1). This RNA fraction was shown to contain histone mRNA by *in vitro* translation. When the wheat embryo cell-free system¹² was programmed with 7–11S RNA, the five major histones were detected as products. Before Poly(U)–Sepharose chromatography of the RNA, additional products were also detected (data not shown). However, up to 90% of the RNA isolated from the second cycle of sucrose gradient fractionation (Fig. 1) passed

through Poly(U)–Sepharose as unbound material. This RNA programmed the synthesis of histones alone as judged by analysis of protein products on two separate acrylamide-gel electrophoresis systems (Fig. 2). Although it is possible that non-histone translation products (if present) would not enter the low pH gels¹⁴ (Fig. 2b), this would not be the case for SDS-urea gels run near neutral pH (ref. 13) (Fig. 2a). For both systems, the products of *in vitro* translation correspond precisely with marker ¹⁴C-histones electrophoresed in the same gel. These results suggest that the poly(A) minus RNA preparation is free of detectable contaminating mRNA species.

Characterisation of histone cDNA

Although the 7–11S RNA preparation programmed the synthesis *in vitro* of histones alone, the presence of other kinds of non-polyadenylated RNA without mRNA activity could not be excluded.

It seemed reasonable to assume that rRNA breakdown products would be the most abundant contaminants of the histone mRNAs. Rather than attempt further purification of the RNA sequences themselves, we synthesised randomly primed cDNA¹⁵ and then removed cDNA complementary to rRNA from the mixture. The ³²P-cDNA preparation was hybridised to a vast excess of 28S and 18S rRNA and the mixture centrifuged

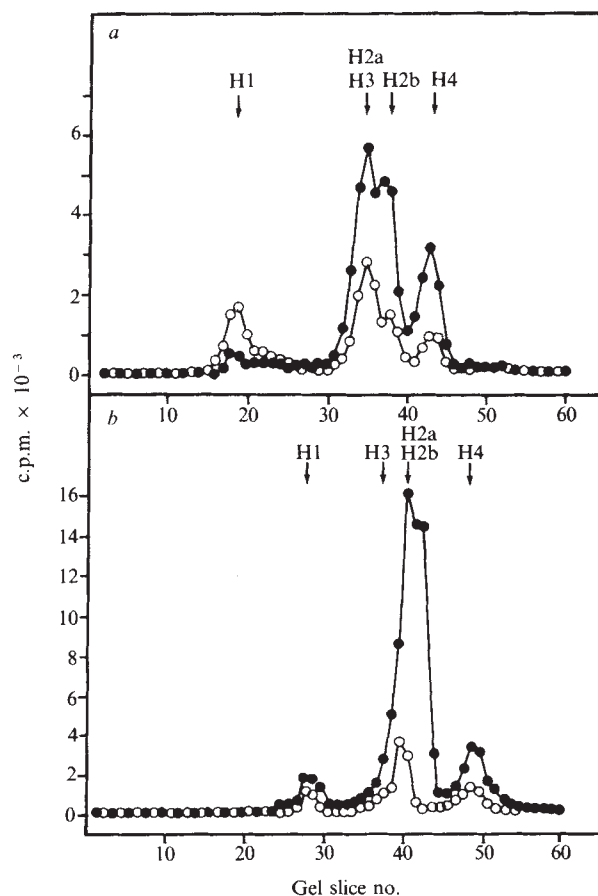


Fig. 2 Analysis on polyacrylamide gels of the *in vitro* translation products of 7–11S RNA unbound to poly(U)–Sepharose. 0.5 μ g of the unbound 7–11S RNA was translated in the wheat-germ cell-free system of Roberts and Paterson¹² at optimal salt conditions of 80 mM K^+ and 3 mM Mg^{2+} . ³H-leucine-labelled translation products were electrophoresed on: a, the SDS-urea gels of Swank and Munkres¹³, and b, the low-pH urea gels of Panyim and Chalkley¹⁴. The gels were then sliced and the radioactivity measured. ●, ³H-labelled product of *in vitro* translation; ○, ¹⁴C-labelled chicken histone standards.

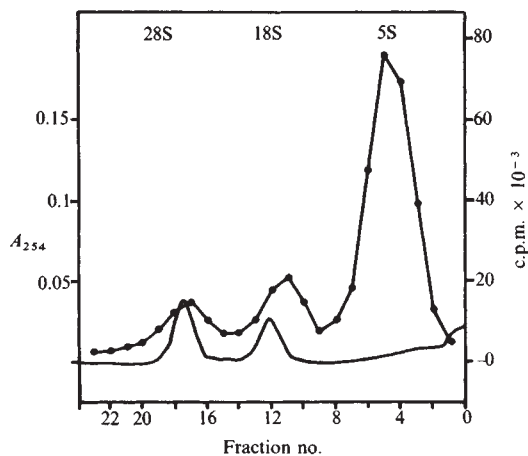


Fig. 3 Preparation of ^{32}P -chicken histone cDNA. DNA complementary to the 7–11S RNA sequences was synthesised with reverse transcriptase using random primers¹⁵. With 2 μg of RNA and ^{32}P -dATP of specific activity 50 Ci mmol^{-1} , we routinely obtained a yield of 1.5×10^7 d.p.m. in cDNA. This product in $2 \times \text{SSC}/0.5\%$ SDS was hybridised to a vast excess of 18S and 28S rRNA at 60°C to $R_0/1$. Ribosomal cDNA sequences were separated from putative histone cDNA sequences by sucrose gradient centrifugation (10–40% sucrose gradients centrifuged at 200,000g, 16 h at 4°C). Fractions of 0.5 ml were collected from the gradient and ^{32}P -radioactivity determined. The cDNA sedimenting at about 5S was pooled and used as a histone gene probe. ●, ^{32}P c.p.m.; —, A_{254} .

on sucrose gradients (Fig. 3). About 30% of the cDNA sedimented with 18S and 28S rRNA, and the remaining 70% of the cDNA was present as a broad peak centred around 5S.

It was possible to test directly whether the 5S cDNA contained histone gene sequences. In a 'Southern' transfer experiment (Fig. 4), the 5S cDNA cross-hybridised to cloned sea urchin histone gene sequences (*Echinus esculentus* histone genes: λ clone 55, K. Murray and K. Peden, unpublished observations).

To test the proposal that this 5S cDNA contained predominantly histone coding sequences, the ^{32}P -labelled cDNA was hybridised to total genome DNA from *E. esculentus*. *EcoRI*-digested sea urchin DNA was electrophoresed on agarose gels and transferred to nitrocellulose filters¹⁶ (Fig. 5). It is possible to detect two closely migrating fragments (about 6 and 7 kilobases) which are identical in size to those detected by homologous *E. esculentus* histone gene probe. No other cross-hybridisation bands were detected when the chicken histone cDNA probe was used. (There are two separate size classes of histone gene repeats within the *E. esculentus* genome. λ Clone 55 contains the larger (7 kilobase) histone gene repeat unit (K. Murray: λ clone 55, K. Murray and K. Peden, unpublished observations).)

As discussed above, the most likely contaminants in the chick histone cDNA probe are ribosomal cDNA sequences. When cDNA probes from chicken rRNA were cross-hybridised to *EcoRI*-digested *E. esculentus* DNA (transferred to filters), the bands observed were readily distinguished from the characteristic 6 and 7 kilobase bands containing histone genes (Fig. 5c).

Therefore, we have shown that the 7–11S RNA prepared from chick embryos and passaged through poly(U)–Sepharose, programmes the synthesis of histones alone in an *in vitro* translation system. When cDNA is prepared from this RNA, and sequences complementary to rRNA are removed, the remaining cDNA can be used to detect sea urchin histone gene sequences. We have used this probe to detect histone genes in chicken DNA.

The reiteration frequency of chicken histone genes

When chicken histone cDNA sequences were reannealed with genomic DNA from adult chickens, a single transition was

observed with a $C_0t_{1/2}$ value of 10^2 (Fig. 6). This transition is distinct from that of unique sequences which have a $C_0t_{1/2}$ value of 10^3 . These data show that histone genes are reiterated about 10-fold in the adult genome.

The rate of reannealing of histone genes is also quite distinct from that of ribosomal genes ($C_0t_{1/2}$ value of 10 Fig. 6), and the data show that the chicken histone cDNA contained negligible amounts of ribosomal sequences.

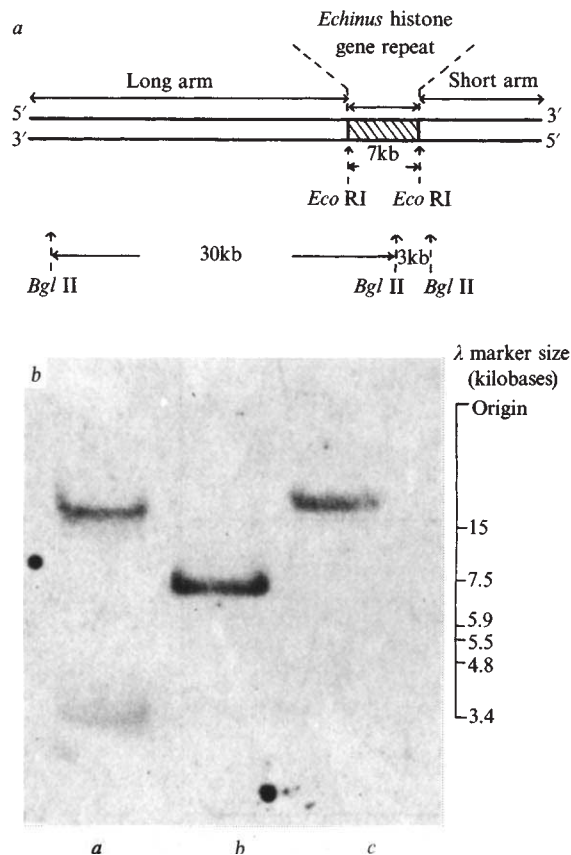


Fig. 4 Identification of chicken histone cDNA sequences, using cloned sea urchin histone genes. DNA from λ clone 55 (containing the 7-kilobase histone gene repeat from *E. esculentus*) was digested with *EcoRI* or *BglII*. Samples of 0.5 μg were electrophoresed on 1% agarose gels (14 cm long, 2 mm thick, 1 cm lane width) at 60 mA for 4 h. DNA fragments were transferred to nitrocellulose¹⁶, baked in a vacuum oven for 3 h at 80°C , and soaked in Denhardt's solution¹⁷ in $2 \times \text{SSC}$ for 6 h at 65°C . Filters were dipped in ^{32}P -labelled chicken histone probe (prepared as described in Fig. 3) in $2 \times \text{SSC}/0.5\%$ SDS (10^6 d.p.m. per μg cDNA). Strips were overlaid with paraffin oil and hybridisation allowed to proceed for 16 h at 65°C . After extensive washing in $2 \times \text{SSC}/0.5\%$ SDS at 65°C , dry filters were put in contact with X-ray film for 2 days, and autoradiograms developed. The size of fragments were determined by reference to λ markers (cleaved with *HindIII* or *EcoRI*) run in the same gel, and detected by ethidium bromide staining. a, *EcoRI* and *BglII* restriction sites within λ clone 55. *EcoRI* excises the complete 7-kilobase histone gene insert and *BglII* cleaves the insert into two fragments. The sizes shown for the fragments are approximate. b, λ clone 55 restriction fragments hybridised to chicken histone cDNA and detected by autoradiography. a, λ clone 55 DNA restricted with *BglII*; b, λ clone 55 DNA restricted with *EcoRI*; c, unrestricted λ clone 55 DNA.

Restriction endonuclease analysis of chicken histone genes

We have not been able to detect histone gene sequences in chicken DNA using a cloned sea urchin histone gene probe (K. Murray, K. Peden and J.R.E.W.). However, as shown in Fig. 5, the chicken cDNA probe can be used to detect *EcoRI* fragments

in sea urchin DNA of the expected size for the histone gene repeat. The reason for this apparent anomaly is probably related to gene numbers. Several hundred histone genes in sea urchin genomic DNA¹⁸ could be detected by the chicken histone probe, but the 10 histone genes in the chicken genome were not detectable with the sea urchin probe. For all subsequent experiments involving restricted chicken DNA, we have used the homologous chicken probe.

Adult genomic DNA was prepared from chicken red blood cells²⁰ cleaved with the appropriate enzyme, and electrophoresed on agarose gels. After transfer of the DNA to nitrocellulose, the filters were impregnated with a solution containing ³²P-histone cDNA, and hybridisation was allowed to proceed at 65 °C for 16 h. Specific sequences were detected by autoradiography. Because of the sensitivity of the transfer procedure¹⁶, reiterated sequences such as ribosomal genes (100-fold reiterated in chicken, Fig. 6) are readily detected. We considered it necessary to determine the position of ribosomal gene fragments in endonuclease-restricted chicken DNA. In the event of trace contamination of histone cDNA with ribosomal cDNA sequences, ribosomal gene fragments could then be identified.

Tracks *b*, *d*, *f*, *h* in Fig. 7 show the location of ribosomal gene sequences on the same nitrocellulose filter previously used to detect histone genes. To achieve this, the hybridised ³²P-histone cDNA was allowed to decay sufficiently before the ribosomal cDNA probe was hybridised. Direct comparison could thus be made between the size of fragments containing the two different sets of genes within one transfer. For each of the four restriction enzyme experiments, histone genes are in different sized fragments of DNA compared with ribosomal genes.

Chicken histone cDNA was used to detect histone gene sequences in restriction endonuclease-cleaved chicken DNA. Fragments containing histone genes were detected in chicken DNA cleaved with *Bgl*II, *Eco*RI, *Bam*HI and *Hind*III (Fig. 7,

tracks *a*, *c*, *e* and *g*, respectively in Fig. 7). One outstanding feature in tracks *a*, *c* and *e* is the presence of a fragment about 15 kilobases long. We propose that this is the repeat length for the histone genes in the chicken genome.

Histone fragments generated by *Hind*III are shown in Fig. 7*g*. The additive size of the three *Hind*III histone DNA fragments (13–14 kilobases) is marginally less than the 15-kilobase repeat length seen consistently with *Eco*RI, *Bgl*II and *Bam*HI. One possible explanation is that a fourth histone fragment (between 1 and 2 kilobases long) is generated by *Hind*III cleavage. If the majority of this additional fragment was spacer DNA (sequences not represented in histone mRNAs), then the fragment would not be detected by the chicken histone gene probe.

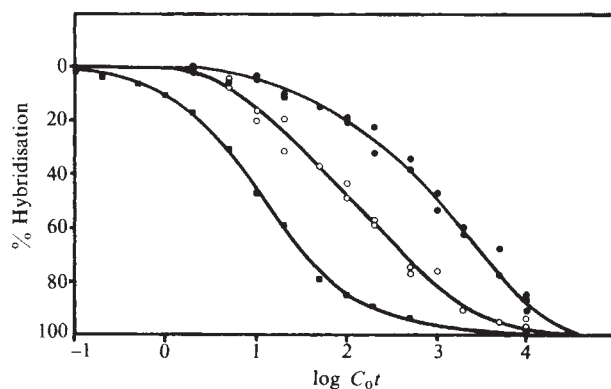


Fig. 6 Estimation of histone gene reiteration frequency in the chicken genome. Reassociation reactions of ³H-labelled total chicken DNA (●) and of unlabelled genomic DNA in the presence of ³²P-labelled histone cDNA (○) or ribosomal cDNA (■) were set up and samples taken as described by Kemp¹⁹, except that 0.18 M NaCl was used in the hybridisation buffer. The degree of reassociation was assayed using nuclease S₁. At the highest *C*₀*t* values, the maximum level of hybridisation was about 70%; the data have been normalised to 100%.

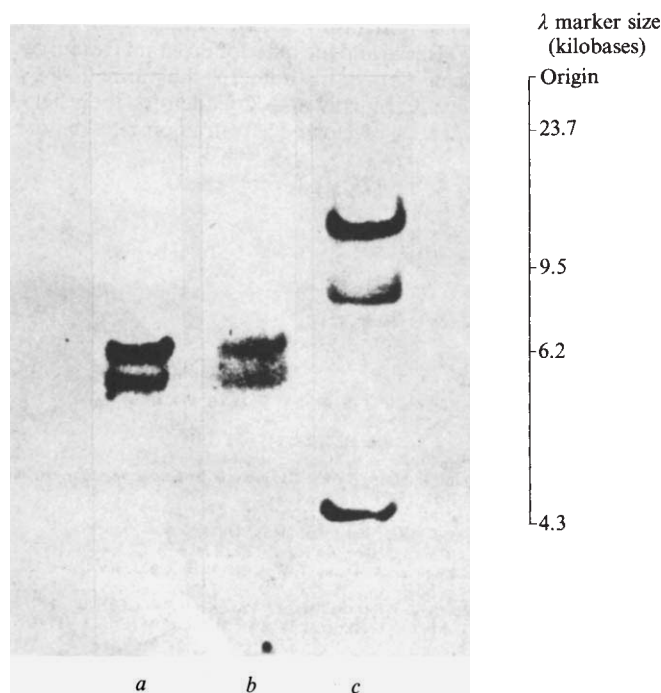


Fig. 5 Detection of 6- and 7-kilobase bands in sea urchin DNA, with chicken histone cDNA. Sea urchin DNA was digested to completion with *Eco*RI. Samples of 2 µg were electrophoresed on 1% agarose gels (30 cm long, 1 cm thick, 1 cm lane width) at 60 mA for 16 h. Fragments containing sea urchin histone genes were detected as described in Fig. 4. *a*, Bands (about 6 and 7 kilobases) detected with ³²P-*E. esculentus* histone gene probe. *b*, Bands in sea urchin DNA detected with chicken histone cDNA; *c*, Bands in sea urchin DNA detected with cDNA made on 18S and 28S chicken rRNA templates.

It is also apparent that both *Bgl*II and *Bam*HI digests reveal sequence heterogeneity in the histone genes. We have seen this heterogeneity in DNA prepared from single animals. In the case of the *Bam*HI histone gene pattern (Fig. 7*e*), two lower molecular weight bands are visible in addition to the 15-kilobase fragment. These two smaller bands (about 11.5 and 4.5 kilobases, respectively) together approximate to the size of the 15-kilobase repeat unit. They are not due to partial *Bam*HI cleavage, as there was no change in the histone gene pattern following further enzymatic digestion. This evidence suggests that there are two classes of histone gene repeats with respect to *Bam*HI cleavage sites; that is, one class contains one and the other class two restriction sites per histone gene repeat unit.

Two classes of histone gene sequences are also detected when chicken DNA is cleaved with *Bgl*II (Fig. 7*a*). One class is cleaved once per repeat to generate a 15-kilobase fragment; the other class does not contain restriction sites for *Bgl*II, as the high molecular weight DNA containing histone gene sequences hardly enters the gel. This material was not susceptible to further digestion with *Bgl*II. It is possible to separate the high molecular weight *Bgl*II histone gene fragment from the 15-kilobase *Bgl*II fragment by sucrose gradient centrifugation (data not shown). As indicated in Fig. 8 (track *c*), digestion of this high molecular weight *Bgl*II fragment with *Eco*RI gives rise to the unit histone gene repeat of 15 kilobases in the same way as does total chicken DNA when digested with *Eco*RI (Fig. 8*a*).

It is evident that there is some heterogeneity in the histone genes, and further, that precise details of chicken histone gene organisation require cloned fragments. However, given that restriction enzyme patterns show (1) a 15-kilobase histone unit alone, (2) a higher molecular weight histone fragment which can generate a 15-kilobase histone unit, or (3) histone gene bands

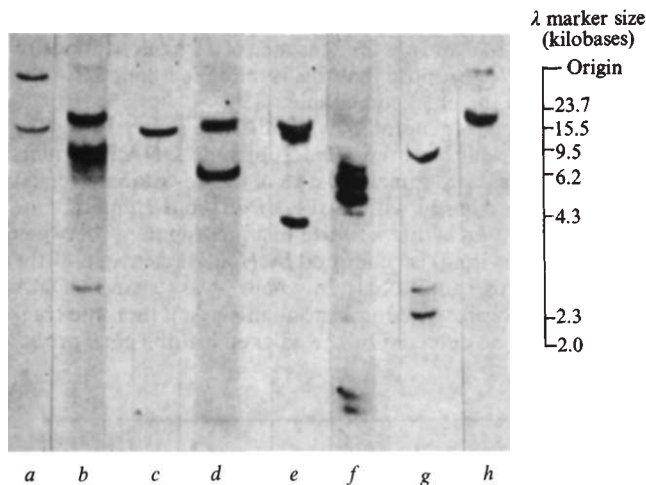


Fig. 7 Detection of histone genes in chicken DNA cleaved with restriction enzymes. DNA isolated from chicken erythrocytes²⁰ was digested with restriction enzymes *EcoRI*, *BglII*, *BamHI* or *HindIII*. After phenol extraction and ethanol precipitation, the digested DNA (50 µg per track) was electrophoresed on 1% agarose gels, transferred to nitrocellulose¹⁶ and hybridised with ³²P-labelled chicken histone cDNA, prepared as described in the legend to Fig. 3. After the histone fragments had been detected by autoradiography, the nitrocellulose strips were stored for 4 weeks to allow a significant level of decay of the ³²P-histone probe. The same nitrocellulose strips were then hybridised with ³²P-labelled cDNA synthesised from 18S and 28S chicken ribosomal RNA templates. The size of the fragments was determined using *HindIII*-cleaved λDNA and *Sall*-cleaved λDNA electrophoresed in the same gel as a standard. Lanes, a, c, e and g were hybridised with chicken histone cDNA, and they contain chicken DNA digested with *BglII*, *EcoRI*, *BamHI* and *HindIII*, respectively. Lanes b, d, f and h were hybridised with chicken 18S and 28S ribosomal cDNAs, and they contain chicken DNA digested with *BglII*, *EcoRI*, *BamHI* and *HindIII*, respectively.

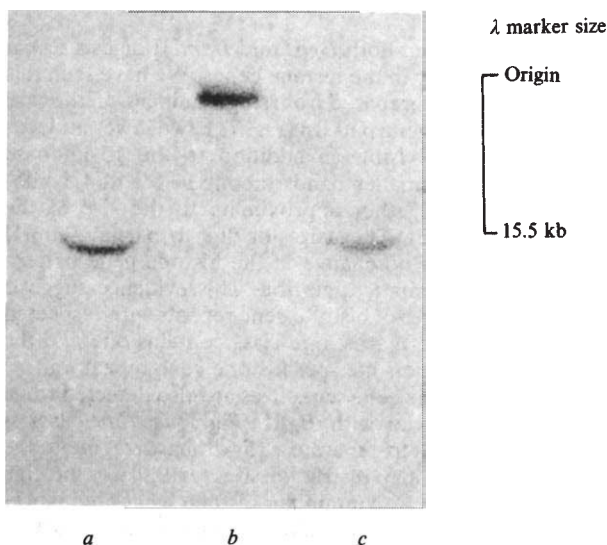


Fig. 8 Digestion of the high molecular weight *BglII* histone gene fragment with *EcoRI*. Chicken DNA was digested to completion with *BglII* and centrifuged on 10–40% sucrose gradients in 0.01 M NaCl, 0.001 M EDTA, 0.01 M Tris-Cl, pH 7.5, at 200,000g at 4 °C for 16 h in a Beckman ultracentrifuge. The large *BglII* histone gene fragment was separated from the 15-kilobase *BglII* fragment (Fig. 7a) on fractionation of the gradient (data not shown). After *EcoRI* digestion of the DNA fraction containing the large *BglII* histone gene fragment, the DNA was electrophoresed on 1% agarose, transferred to nitrocellulose and hybridised with ³²P-chicken histone cDNA as described in the legend to Fig. 4. a, Histone gene sequences in total chicken DNA digested with *EcoRI*; b, large *BglII* chicken histone gene fragments fractionated by sucrose gradient centrifugation; c, large *BglII* histone gene fragment in chicken DNA fractionated by sucrose gradient centrifugation and digested with *EcoRI*.

which add up to about 15 kilobases in size, we conclude that chicken histone genes are clustered and have a unit repeat length of 15 kilobases.

We believe this to be the first direct evidence for the histone gene repeat size in a higher eukaryote. It is not clear whether the 15-kilobase repeat contains genes coding for each of the five classes of histone proteins. The only histone gene repeats so far studied are those from sea urchin^{1–4} and *Drosophila*⁵, and in these organisms at least, the histone repeat contains a complete set of histone genes.

More specifically, we have shown previously that the H5 genes (only expressed in erythroid cells) are reiterated 10-fold in the chicken genome²¹, and we show here that other histone genes are reiterated to the same extent. It is not clear whether H5 genes are contained within the 15-kilobase repeat. Nevertheless, there seems to be at least four to five times more DNA per repeat than is required for coding. It will be interesting to determine whether histone mRNAs are transcribed from one DNA strand (as in sea urchin³) or from both strands (as in *Drosophila*⁵). The report of a 15-kilobase RNA species containing histone sequences in HeLa cells²² suggests the former possibility.

If 15 kilobases is representative of higher eukaryote histone gene repeat size, then this is about 2.5–3 times the length of its counterpart in sea urchin DNA and *Drosophila* DNA. If the arrangement of coding sequences and A/T-rich spacer sequences found in sea urchin histone genes⁴ is similar in higher eukaryote histone genes, there would be a significant increase in length of 'spacer' DNA, as the length of coding regions for histones will be almost identical in all species. Nevertheless, the length of spacer DNA in a cistron cannot necessarily be equated with the evolutionary complexity of the organism. For example, the repeat length of ribosomal genes in *Dictyostelium* and mouse is comparable (38 (ref. 23) and 44 (ref. 24) kilobases, respectively) and both of these are much larger than the 11–14-kilobase repeat in *Xenopus*²⁵. Whether the extra DNA in higher eukaryote histone genes exists as intervening sequence DNA within coding regions or as elongated spacer regions between histone coding regions, remains to be determined.

We thank Dr J. W. Beard and the office of Program Resources and Logistics, National Cancer Institute, for the purified AMV polymerase, R. Saint, K. Murray and K. Peden for their help, and K. Gross for details of clone 55 restriction sites before publication.

Received 17 July 1978; accepted 19 March 1979.

- Gross, K., Schaffner, W., Telford, J. & Birnstiel, M. L. *Cell* **8**, 479–484 (1976).
- Kedes, L. H. *Cell* **8**, 321–331 (1976).
- Cohn, R. H., Lowry, J. C. & Kedes, L. H. *Cell* **9**, 147–161 (1976).
- Portmann, R., Schaffner, W. & Birnstiel, M. L. *Nature* **264**, 31–34 (1976).
- Lifton, R. P., Goldberg, M. L., Karp, R. W. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1047–1052 (1977).
- Burckhardt, J. & Birnstiel, M. L. *J. molec. Biol.* **118**, 61–79 (1978).
- Wilson, M. C. & Melli, M. J. *molec. Biol.* **110**, 511–535 (1977).
- Bos, E., Roskain, W. & Gallwitz, D. *Biochem. biophys. Res. Commun.* **73**, 404–410 (1976).
- Borun, T. W., Ajiro, K., Zweidler, A., Dolby, T. W. & Stephens, R. E. *J. biol. Chem.* **252**, 173–180 (1977).
- Destrée, O. H. J., Haenni, A.-L. & Birnstiel, M. L. *Nucleic Acids Res.* **4**, 801–811 (1977).
- Seeburg, P. H., Shine, J., Martial, J., Baxter, J. D. & Goodman, H. M. *Cell* **12**, 157–165 (1977).
- Roberts, B. E. & Paterson, B. M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2330–2334 (1973).
- Swank, R. T. & Munkres, K. D. *Analyt. Biochem.* **39**, 462–477 (1971).
- Panyim, S. & Chalkley, R. *Biochemistry* **8**, 3972–3979 (1969).
- Melli, M., Spinnelli, G., Wyssling, H. & Arnold, E. *Cell* **11**, 651–661 (1977).
- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
- Denhardt, D. *Biochem. biophys. Res. Commun.* **23**, 641–646 (1966).
- Kedes, L. H. & Birnstiel, M. L. *Nature new Biol.* **230**, 165–169 (1971).
- Kemp, D. J. *Nature* **254**, 573–577 (1975).
- Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36**, 32–38 (1973).
- Scott, A. C. & Wells, J. R. E. *Nature* **259**, 635–638 (1976).
- Melli, M., Spinnelli, G., Wyssling, H. & Arnold, E. *Cell* **11**, 651–661 (1977).
- Maizels, N. *Cell* **9**, 431–438 (1976).
- Cory, S. & Adams, J. M. *Cell* **11**, 795–805 (1977).
- Wellauer, P. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 2823–2827 (1974).

letters

Production of new cosmological perturbations during the radiation-dominated era

PURCELL (unpublished data) has noted the tendency for 'dust' particles strongly to cluster when they are coupled by viscous drag to a 'gas' whose velocity field is stochastic:

$$\frac{dV_{\text{dust}}}{dt} = \alpha \{V_{\text{gas}}(x, t) - V_{\text{dust}}\} \quad (1)$$

where α (s^{-1}) is the viscous coupling constant. The application of the Purcell clustering effect to cosmology which is considered here generates relatively large isothermal perturbations from relatively small adiabatic perturbations which will not otherwise survive damping during recombination.

The Purcell clustering effect seems at first to be paradoxical as one would expect a stochastic gas flow to randomise, rather than cluster, the positions of independent, non-interacting dust particles. The explanation is that the velocity field is not completely random but consists of disturbances that propagate at sound velocity. Therefore, on a given scale the gas velocity will be correlated for one sound travel time. In linear order in the gas velocity the dust particles always return to their original positions after the passage of an acoustic wave; in quadratic order, however, transport effects are present^{1,2}. The accelera-

tion in quadratic order on a given scale will go as V_{gas}^2/l . In the low frequency limit, that is, for scales $l \gg c_s \alpha^{-1}$, particles are strongly coupled (frozen in) to the gas. The particles will therefore acquire a velocity relative to the gas of $u \sim V_{\text{gas}}^2/\alpha l$. The mean square clustering of particles due to this anomalous motion per sound travel time is $\delta^2 \sim (u/c_s)^2$. The clustering of the particles at later times will not be correlated with the original clustering and will add randomly in the mean square sense. The growth of particle clustering with time will therefore be

$$\left(\frac{\delta\rho}{\rho}\right)_t^2 \sim \left(\frac{V_{\text{gas}}}{c_s}\right)^4 \left(\frac{c_s}{l\alpha}\right)^3 \alpha t \quad (2)$$

for $l \gg c_s \alpha^{-1}$. On small scales $l \ll c_s \alpha^{-1}$, the particles will always be in the process of relaxing towards the velocity u given above. The typical response to the quadratic term will be $(1 - e^{-\alpha/\omega})u$ where ω is the frequency corresponding to the scale l . The anomalous motion of the particles relative to the gas will therefore be $V_{\text{gas}}^2(\alpha/\omega^2)$. In the same manner we may see that in the high frequency limit the particle clustering is given by

$$\left(\frac{\delta\rho}{\rho}\right)_t^2 \sim \left(\frac{V_{\text{gas}}}{c_s}\right)^4 \left(\frac{\alpha}{\omega}\right) \alpha t \quad (3)$$

We will give a detailed analytic derivation of the above equations elsewhere. Without considering the above physical picture the ensemble-average mean square increase of clustering can be computed using time from a formal expansion of equation (1)

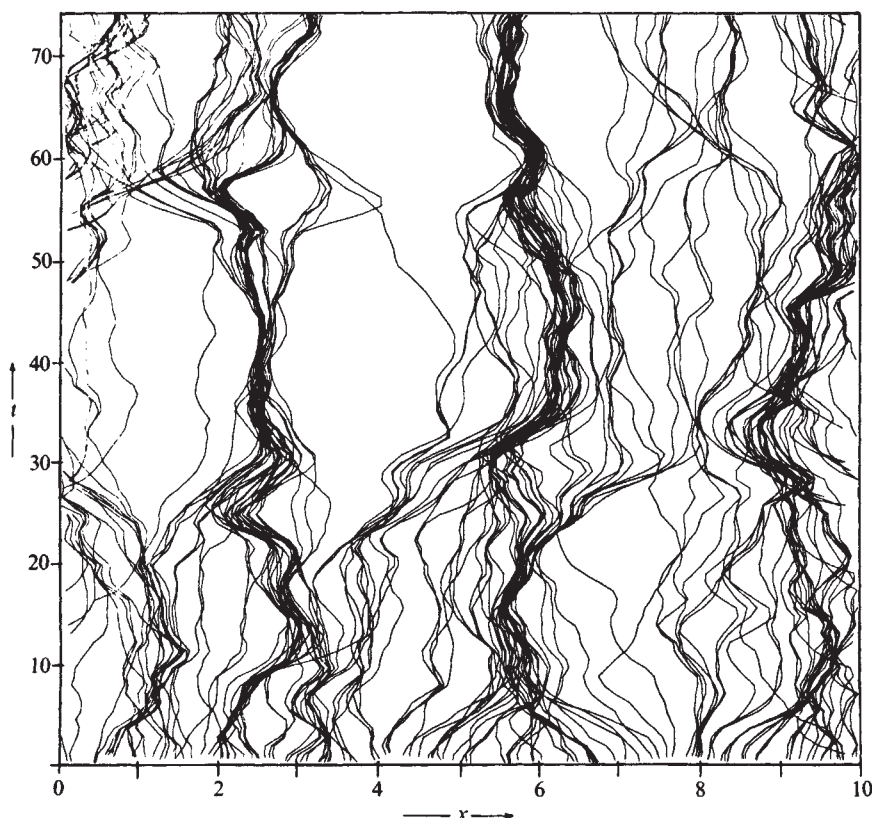


Fig. 1 Space-time diagram showing the aggregation of particles by a random acoustic field. First order motion, which integrates to zero over time, has been removed from the picture. Gaussian acoustic white noise propagates both leftwards and rightwards at velocity $dx/dt = 1$. The viscous coupling time of particles to the gas is $\alpha^{-1} \approx 2$ t -units, so clustering is primarily on a scale of about 2 x -units. In the period shown, about 40 uncorrelated small consolidations add in the mean-square sense to give a density perturbation of order unity.

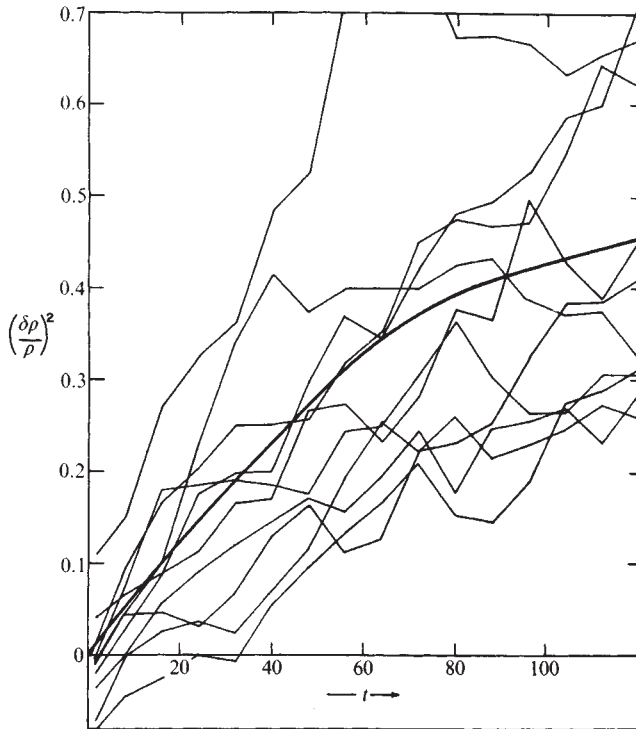


Fig. 2 Results from many runs as in Fig. 1. The squared clustering measure $(\delta\rho/\rho)^2$ evaluated at a spatial scale of 2 x -units is shown plotted against time. Narrow lines are 10 individual runs; the heavy line is the mean of 30 further runs. For $(\delta\rho/\rho)^2 \leq 0.3$, squared clustering varies linearly with time, as predicted by the analytic theory. For larger $(\delta\rho/\rho)^2$ deviations from linear growth become visible, due to numerical brownian motion on the differencing scale.

and the equation of continuity. Here we summarise the results of numerical experiments which corroborate the analytic derivations, and with an immediate application to cosmology.

An example of a numerical experiment in one dimension is shown in Fig. 1, where 128 particles are shown clumping to order unity density contrast on a scale containing about 25 particles. In this experiment density contrast was measured at the scale where $\omega \sim \alpha$. (Only the scale, not the number of particles shown, is relevant: 12,800 particles in the same hydrodynamic velocity field would have clustered on scales containing 2500 particles.)

Figure 2 shows $\delta\rho/\rho$ as a function of time for 10 runs which differ only in having different random sequences of sound waves. Also shown is the mean of 30 additional runs. The secular nature of the clustering effect is evident. In equation (2), the value of the coefficient taken from the experiments agrees with the value computed analytically to within 20%. In three dimensions, equation (2) also holds, with a somewhat different numerical coefficient. The coefficient is weakly dependent on the slope of the power spectrum of the adiabatic perturbations.

We now consider the application of the Purcell clustering effect to cosmology, namely its ability to generate relatively large isothermal perturbations (which can make globular clusters and perhaps galaxies in the post-recombination era) from relatively small adiabatic perturbations which will not otherwise survive damping during recombination. Purcell clustering is able to generate these perturbations on scales much smaller than the Jeans mass, during the radiation dominated era ($z > 10^3$) when, according to current views, matter and radiation are tightly coupled.

Weinberg³ and others have recently pointed out that baryon non-conserving reactions at high energies (a consequence of the

charge-conjugation parity (CP) violation in super-unified theories) will 'cook' the early Universe to a unique value of entropy per baryon, one which in particular cannot vary from point to point spatially. Therefore late-time processes like Purcell clustering (or shock dissipation, see below) may well be the only way to introduce isothermal (entropy) perturbations into a cosmology.

In our cosmological application, the dust particles of Purcell clustering are individual protons and helium nuclei of the pre-recombination, radiation dominated, Friedmann cosmology; the dominant component of the gas is photons; the ionised plasma (dynamically insignificant) supplies the necessary scattering opacity to render the mean free path of photons much shorter than any other scale of interest. However, the finite mean free path of the photons will cause acoustic waves to damp long before the time when $k = \alpha c_s^{-1}$ [ref. 4].

The baryons are viscously coupled to the photon gas primarily by radiation drag on their attendant electrons⁵, with

$$\alpha = \frac{4\sigma_T a T^4}{3m_p c} = 3.1 \times 10^{-11} \text{ s}^{-1} \left(\frac{z}{1300}\right)^4 \quad (4)$$

where σ_T is the Thompson cross section, m_p the proton mass, and z is the redshift of the epoch considered. The temperature now is taken to be 3 K. The Hubble time t_h as a function of z , when the Universe was photon-radiation dominated⁶, is

$$t_h = a/\dot{a} = \left(\frac{8\pi}{3} G\rho\right)^{-1/2} = 3.0 \times 10^{13} \text{ s} \left(\frac{1300}{z}\right)^2 \quad (5)$$

So,

$$\alpha t_h = 940(z/1300)^2 \quad (6)$$

The sound speed in the radiation dominated era is $c_s = c/3^{1/2}$, so the characteristic baryon rest mass contained in a volume of characteristic size c_s/α is

$$M_p = \frac{4\pi}{3} \rho_m \left(\frac{\pi c_s}{\alpha}\right)^3 = 1.7 \times 10^{10} M_\odot \left(\frac{1300}{z}\right)^9 \left(\frac{\Omega}{0.1}\right) \left(\frac{H_0}{60}\right)^2 \quad (7)$$

If we substitute equations (7) and (6) into equation (2) we obtain

$$\left(\frac{\delta\rho}{\rho}\right)_{\text{iso}} \sim 310 \left(\frac{\delta\rho}{\rho}\right)_{\text{ad}} \left[\left(\frac{\Omega}{0.1}\right) \left(\frac{H_0}{60}\right)^2\right]^{1/2} \quad (8)$$

where we have taken a coefficient of 10 in equation (2) from more exact calculations. Equation (8) is valid only for $M > M_c$, the damping scale at recombination. We note that $M_c = 9.8 \times 10^{12} [(\Omega/0.1)(H_0/60)^2]^{1/2} M_\odot$ [ref. 7]. For $M < M_c$ the waves driving the clustering are damped before recombination. As $(M/M_c) = (t/t_h)^{9/4}$ gives the damping time for mass scales $< M_c$ we get

$$\left(\frac{\delta\rho}{\rho}\right)_{\text{iso}} \sim 2 \left(\frac{\delta\rho}{\rho}\right)_{\text{ad}} \left(\frac{M}{M_p}\right)^{5/18} \left[\left(\frac{\Omega}{0.1}\right) \left(\frac{H_0}{60}\right)^2\right]^{7/6} \quad (9)$$

Equations (8) and (9) say, roughly, that 10% adiabatic perturbations will generate 1% isothermal perturbations anywhere in the mass range 10^9 – $10^{12} M_\odot$.

In the 'pancake' theory of galaxy formation as originally proposed by Zel'dovich, Sunyaev *et al.*^{8–11}, the density contrast before recombination is the same at all scales with a value between 10^{-3} and 10^{-4} . In this case Purcell clustering is insignificant. However, calculations by Silk¹² and by Peebles and Uy¹³ seem to show that a considerably larger density contrast on the scale of the pancakes is necessary to produce galaxies, say 10^{-2} . As limits of $< 10^{-3}$ on the large-scale inhomogeneity can be set from the microwave background¹⁴, it seems likely that a viable perturbation spectrum must have $\delta\rho/\rho$ increasing towards smaller masses and reaching the 10% level well within the mass range where Purcell clustering becomes important. In that case the condensation of sub-galactic units renders the subsequent collapse of the pancake stellar dynamical rather than hydrodynamical. The formation of dense, cold sheets needed for fragmentation into galaxies is impossible, and the whole pancake hypothesis must be reexamined.

The Purcell clustering process is physically distinct from another process which can generate entropy fluctuations, namely the dissipation of shock waves that are formed by wave steepening, and wave-wave interactions. For 10% adiabatic fluctuations described above, the Purcell clustering seems to dominate for two reasons: first incoherent acoustical disturbances steepen much more slowly than plane waves¹⁵, and second shock dissipation goes as the cube of shock strength, so weak (10%) shocks are not very dissipative.

This work was supported in part by NSF grant PHY78-09616.

WILLIAM H. PRESS
ETHAN T. VISHNIAC

Harvard-Smithsonian Center for Astrophysics,
60 Garden St.,
Cambridge, Massachusetts 02138

Received 7 February; accepted 20 March 1979.

- Westervelt, P. J. *J. acoust. Soc. Am.* **22**, 319 (1950).
- Ingard, U. & Galehouse, D. C. *Am. J. Phys.* **39**, 811 (1971).
- Weinberg, S. *Phys. Rev. Lett.* (in the press).
- Silk, J. *Nature* **215**, 1115 (1967).
- Peebles, P. J. E. *Physical Cosmology*, 220 (Princeton University Press, 1971).
- Weinberg, S. *Gravitation and Cosmology*, 538 (Wiley, New York, 1972).
- Weinberg, S. *Astrophys. J.* **168**, 175 (1971).
- Zel'dovich, Ya. B. *Astr. Astrophys.* **5**, 84 (1970).
- Sunyaev, R. A. & Zel'dovich, Ya. B. *Astr. Astrophys.* **20**, 189 (1972).
- Doroshkevich, A. G. & Shandarin, S. F. *Astr. Zh.* **51**, 41 (1974); *Sov. Astr.* **18**, 24 (1974).
- Doroshkevich, A. G. & Shandarin, S. F. *Mon. Not. R. astr. Soc.* **182**, 27 (1978).
- Silk, J. *IAU Symp.* **63**, 175 (1974).
- Peebles, P. J. E. & Uy, J. T. *Astrophys. J.* **162**, 816 (1970).
- Smoot, G. F., Gorenstein, M. V. & Muller, R. A. *Phys. Rev. Lett.* **39**, 898 (1977).
- Peebles, P. J. E. *Phys. Rev. D* **1**, 397 (1970).

Evidence for X-ray emission from Kepler's supernova remnant

THE remnant of Kepler's supernova of 1604 AD has long eluded detection at X-ray wavelengths, principally because of its proximity to a source confused region in the direction of the galactic centre. We present here the first evidence for weak soft X-ray emission from Kepler's remnant, which was beyond the capability of earlier X-ray surveys¹ but is now possible because of the increased sensitivity provided by the low energy detectors (LEDs) of the A 2 experiment on the HEAO 1 spacecraft. (The A 2 experiment on HEAO 1 is a collaborative effort led by E. Boldt of GSFC and G. Garmire of CIT, with collaborators at GSFC, CIT, JPL, and UCB). The LEDs are described in detail by Rothschild *et al.*² Briefly, the data reported here were acquired using the LED1 narrow field of view (1.5° FWHM in the scan direction) which has an effective area of 175 cm² and is sensitive to X-rays between 0.2 and 3 keV.

The region of sky containing Kepler's remnant was scanned by LED1 between 11 and 16 March 1978, and Fig. 1 shows the results of superposing five scans during this time interval. The count-rate data are complex owing to the concentration of X-ray sources in the direction of the galactic centre. However, a new soft source, denoted H1735-21, is evident in the data as a count-rate peak centred at a scan angle of 178.2°. This is just resolved from the GX9+9 region. The solid line in Fig. 1 indicates a least-squares fit of the LED1 angular response function to the data bounded by the sources GX9+9 and GX3+1. The 90% confidence error box derived from this fit is depicted in Fig. 2, together with the positions of all known nearby X-ray sources. Incorrect modelling of the background could introduce small systematic error (~0.1°) in position of narrow direction of the error box. Kepler's remnant lies within our error box, and is therefore proposed as the probable identification. We have checked other classes of object which could produce soft X-ray emission, but no other compelling counterparts were found within the error box. The Uhuru source 4U1730-22 (ref. 3), which lies at the edge of the 90% confidence error box, is an unlikely identification because it is

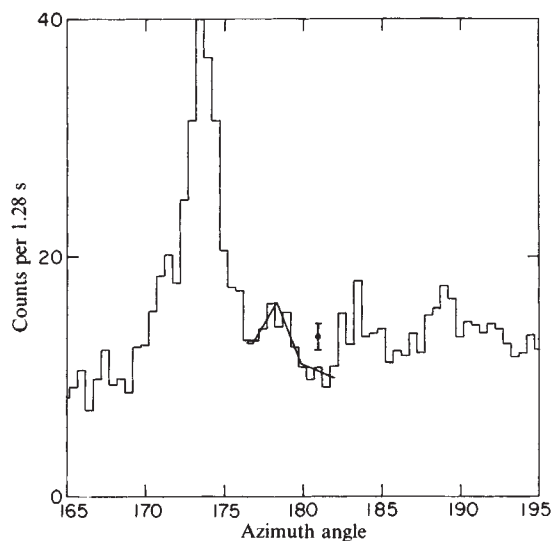


Fig. 1 A superposition of five scans of HEAO A-2 LED data (0.3-3.0 keV). A typical error bar around the region of interest is shown. The solid curve represents the best fit to the triangular response of our detectors to H1735-21 plus the local background. GX9+9 (angle = ~174) and GX3+1 (angle = ~183) constrained the use of background data.

known to be a transient⁴. The spectrum of transients are hard near maximum intensity. The absence of H1735-21 in medium and high energy detectors (S. Pravdo, personal communication) precludes it being the transient detected during an 'on-state'. Our data are consistent with the source being constant during the period of observation. However, because the source is just at our level of detectability, we are insensitive to decreases in intensity.

The distance to Kepler's remnant remains uncertain. Because it lies in the direction of the galactic centre the application of normal distance measuring techniques (for example, neutral hydrogen absorption observations) is impossible. A position within the nuclear bulge, with distance of about 10 kpc, has been invoked to help explain the comparatively large angular displacement from the galactic plane. The historically inferred apparent magnitude of -3, and its classification as Type I (ref. 5), suggest a distance of between 3.5 and 6 kpc, assuming an

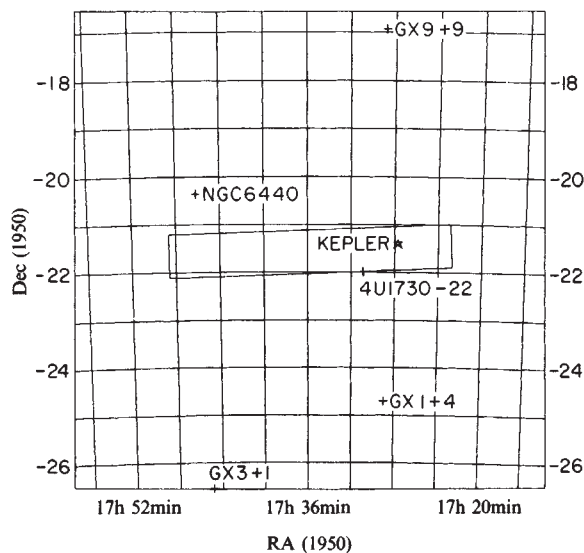


Fig. 2 A map of the region shows the derived HEAO A-2 90% confidence error box. The position of Kepler's SNR as well as all known nearby X-ray sources is shown. Note the position of 4U1730-22 is at the edge of the box.

absolute magnitude for a Type I supernova of -19 and visual absorption in the range 0.5 to 1 mag kpc^{-1} . Recent unpublished observations of the reddening of the optical remnant of Kepler's supernova (I. J. Danziger, personal communication) suggest a distance of $\sim 4 \text{ kpc}$.

The intensity derived for H1735-21 is $(6.3 \pm 1.5) \times 10^{-2}$ counts $\text{cm}^{-2} \text{s}^{-1}$ between 0.3 and 3.0 keV , corresponding to a flux of $(2.4 \pm 0.6) \times 10^{-10} \text{ erg cm}^{-2} \text{s}^{-1}$, assuming a power law spectrum with a photon index of -2.0 and a column density of $10^{22} \text{ H-atoms cm}^{-2}$. (The count flux as well as the energy flux could be uncertain by a factor of 2 due to uncertainties in background subtraction and incident spectrum.) The assumed hydrogen column density, probably appropriate for Kepler's remnant regardless of distance uncertainty, gives an intrinsic X-ray luminosity in the energy range $0.3\text{--}3 \text{ keV}$ of $\sim 180 \times 10^{35} \text{ erg s}^{-1}$ for a distance of 10 kpc , reducing to $\sim 28 \times 10^{35} \text{ erg s}^{-1}$ for a distance of 4 kpc . The former value is greater than for any other known young supernova remnant, including the Crab Nebula, and must cast further doubt on Kepler's remnant lying at the galactic centre distance. The latter value is more acceptable, being comparable with that of other young supernova remnants such as that for Tycho's SNR and Cas A.

On the near side of the Galaxy, six supernovae recorded in historical times have been identified with remnants⁵—those of 185 AD (remnant RCW 86), 1006 AD (PKS 1459-41), 1054 AD (the Crab Nebula), 1181 AD (3C 58), 1572 AD (Tycho) and 1604 AD (Kepler). The objects Cas A, RCW 103 (ref. 6), and MSH 11-54 (refs 7, 8) are also believed to be the remnants of supernovae which have occurred within the past two millennia but were not recorded historically. All the above mentioned objects have previously been detected in X-rays with the exception of Kepler's remnant. For this reason, it has long been classified as a highly likely X-ray source^{9,10}, although problems of confusion and sensitivity have thwarted previous attempts at detection. HEAO 2 will now allow the detection reported here to be confirmed, the source to be imaged, and its spectrum to be more accurately determined.

We thank G. Riegler, P. Agrawal, E. Boldt and the GSFC group for their contributions towards the development of the LED detectors. This work was supported by NASA contract NAS 5-23315.

IAN R. TUOHY

Director's Unit, Research School of Physical Sciences,
Australian National University,
Canberra ACT, Australia

JOHN J. NUGENT

GORDON P. GARMIRE

Department of Physics,
California Institute of Technology,
Pasadena, California 91125

DAVID H. CLARK

Royal Greenwich Observatory,
Herstmonceux Castle,
Hailsham, East Sussex, UK

Received 9 March; accepted 2 April 1979.

1. Burginyon, G. *et al. Astrophys. J.* **179**, 615 (1973).
2. Rothschild, R. *et al. Space Sci. Instrum.* (submitted).
3. Forman, W. *et al. Astrophys. J. Suppl.* **38**, 357 (1978).
4. Cominsky, L., Jones, C., Forman, W. & Tananbaum, H. *Astrophys. J.* **224**, 46 (1978).
5. Clark, D. H. & Stephenson, F. R. *The Historical Supernovae* (Pergamon, Oxford, 1977).
6. Tuohy, I. R. *et al. Astrophys. J. Lett.* (in the press).
7. Goss, M. *et al. Mon. Not. R. astr. Soc.* (in the press).
8. Share, G. *et al. IAU Circ. No.* 3169 (1978).
9. Clark, D. H. & Culhane, J. L. *Mon. Not. R. astr. Soc.* **175**, 573 (1976).
10. Winkler, P. F., Jr *Mem. Soc. Astr. Ital.* **49**, 599 (1978).

A peculiar galaxy system in Hydra

IN the course of scanning the Palomar Sky Survey, I noticed what appeared to be a large and a small elliptical galaxy interconnected by a long straight bridge which could have been a radial jet from the larger galaxy. Furthermore a comparison between red and blue prints showed the smaller galaxy to be

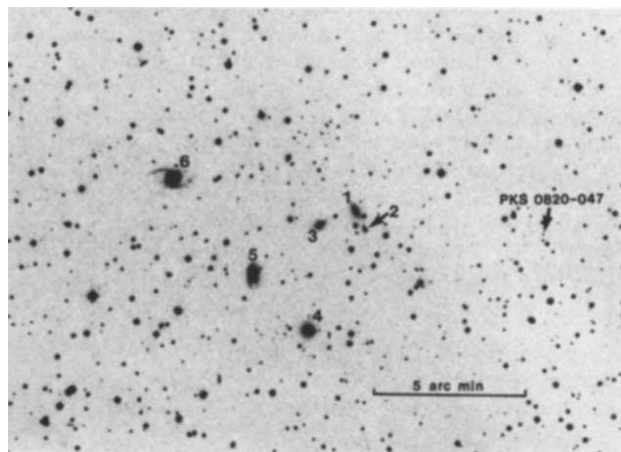


Fig. 1 The cluster of galaxies containing the peculiar system (from the blue print of the Palomar Sky Survey). Identifications of galaxies are as follows: the peculiar system (galaxy 1 = MCG -1-22-6 = F-272; galaxy 2 = F-271; galaxy 3 = MCG -1-22-7 = F-273). Galaxy 4 = NGC2583 = MCG -1-22-8; galaxy 5 = MCG -1-22-9; galaxy 6 = MCG -1-22-10. The optical identification of the radio source PKS0820-047 is also indicated. (Reproduced with permission from the National Geographic Palomar Observatory Sky Survey.)

extremely blue, and thus highly unusual. Detailed enlargement of the images suggest that the larger galaxy may possess chaotic filaments akin to those in M82 or NGC1275. The spectroscopic observations described here show that, as expected, the two galaxies have strong narrow emission lines, but also that a third neighbouring galaxy is involved.

The system concerned appears as part of a medium-poor cluster of galaxies at 1950: 8h 20. 7 min, $-4^{\circ}50'$ in the constellation of Hydra (see Fig. 1). The two larger members of the peculiar system are recorded in Vorontsov-Velyaminov's *Morphological Catalogue of Galaxies*¹ (MCG) although the system is apparently not included in his *Atlas of Interacting Galaxies*². (The designations F-271, F-272 and F-273 refer to the selection of these objects for my current survey of 'compact and bright-nucleus galaxies'³.)

Figure 2 shows further enlargements of the peculiar system. These plates are made from a different copy of the Palomar Sky Survey from that on which the pair of ellipticals were originally noticed and the straight bridge between those galaxies (galaxies 1 and 2) is, in this case, not as conspicuous as it appeared on the red print of the first copy. This seems to be why Vorontsov-Velyaminov² did not pick it up; had he done so it would have made an extreme example of his 'bottle form' category.

The extreme blue colour of galaxy 2 is readily apparent by comparing its images in Fig. 2a and b; similarly, the red filaments of galaxy 1 (the appearance of the outer parts of the galaxy is slightly affected by the mottled sky background of the paper print). Galaxy 3 seems to possess bluish outer regions and condensations.

Spectroscopy was carried out in 1978 November/December with the image tube spectrograph on the 1.9 m Radcliffe Reflector at the South African Astronomical Observatory. Unwidened spectra at 210 Å mm^{-1} covering $3,700\text{--}7,000 \text{ Å}$ with a strong peak in the blue were obtained (and reduced in the same manner as described elsewhere³). The spectrum of galaxy 1 shows strong [O III]5007+4959 as well as narrow H β ; the emission region seems more extensive than the central core of the galaxy (it could be filaments). The uncorrected redshift is found to be $z = 0.0218$ (probable error 0.0005). The spectrum of galaxy 2 reveals strong narrow lines of [O II]3727, H β , [O III]5007+4959 and H α , with $z = 0.0230$. Unexpectedly, galaxy 3 also turned out to have emission lines though they are weaker; it has narrow [O II]3727, H β and [O III]5007+4959 with $z = 0.0226$. NGC2583 (galaxy 4) seems to have a con-

ventional absorption spectrum with $z = 0.0197$ (based on the G-band and H and K lines, with a slight uncertainty in identification). Galaxies 5 and 6 proved to be of low surface brightness and identification of features in their spectra proved very difficult; both may have very weak $H\beta$ emission (as might be expected for such late type spirals) with $z = 0.0240$. (As all six galaxies are relatively faint, the spectrograms show appreciable sky contamination. Thus the overall quality does not permit quantitative measurements of line strengths and so on).

The strong narrow emission lines of galaxies 1, 2 and 3 indicate abnormal nebular emission—as is not uncommon in interacting systems (such as, nos 44–45–46, 156–157, 199, 204 and 210 in my current survey³). The astrophysical interpretation of such is usually controversial. In this case it looks as though galaxy 2 has either been ejected from galaxy 1, or at least has passed through it. If the latter is the case, the tidal turmoil has evoked a sudden wave of star formation providing the excitation mechanism for the nebular emission. But are the red filaments of galaxy 1 merely tidal in origin, and how does galaxy 3 fit into such a picture? The alternative would be in line with ideas favoured by Arp⁴—that galaxy 1 is some sort of active galaxy and galaxy 2 has somehow been thrown out from it—red

filaments would support ejective or explosive processes. The notion of an exploding galaxy immediately brings to mind M82 and one may question whether this is really an explosion or are the active trio merely a manifestation of galaxies drifting through dust? None of these explanations seem satisfactory and further speculation should probably await the photography and spectroscopy of this system by larger telescopes.

Figure 2 also indicates some particularly red and blue stars; it is a little unusual that they should appear to lie so close to the galaxies but there is nothing to suggest that they are not foreground stars. Likewise the radio source PKS0820-047 (= OJ-033 = 4C-4.26) identified with an 18th magnitude starlike object lies close enough to appear in Fig. 1.

I am grateful for the hospitality at the A. J. Dyer Observatory of Vanderbilt University, Nashville, Tennessee (where this object was originally noticed in early 1977) and to the director of the South African Astronomical Observatory for observing time.

A. P. FAIRALL

Department of Astronomy,
University of Cape Town,
7700 Rondebosch, South Africa

Received 21 February; accepted 13 March 1979.

1. Vorontsov-Velyaminov, B. & Arhipova, V. *Morphological Catalogue of Galaxies*, Vol. III (Moscow, 1968).
2. Vorontsov-Velyaminov, B. *Astr. Astrophys. Suppl.* **28**, 1–117 (1977).
3. Fairall, A. P. *Mon. Not. R. astr. Soc.* **180**, 391–400 (1977).
4. Fairall, A. P. *Mon. Not. R. astr. Soc.* (in the press).
5. Arp, H. C. *Astr. Astrophys.* **3**, 418–435 (1969).
6. Solinger, A., Morrison, P. & Markert, T. *Astrophys. J.* **211**, 707–717 (1977).
7. Bolton, J. G. *et al. Austr. J. Phys. Suppl.* **34**, 1–32 (1975).

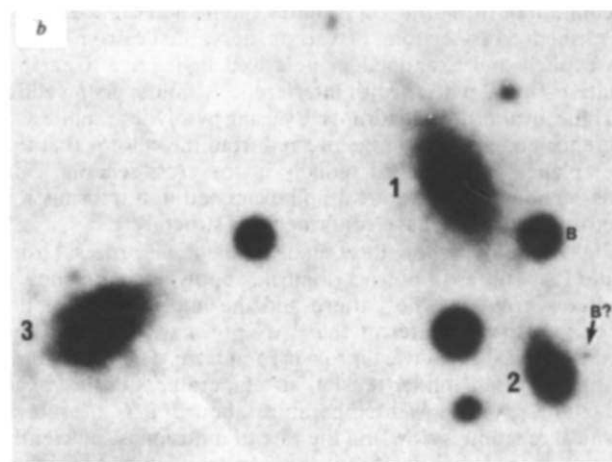
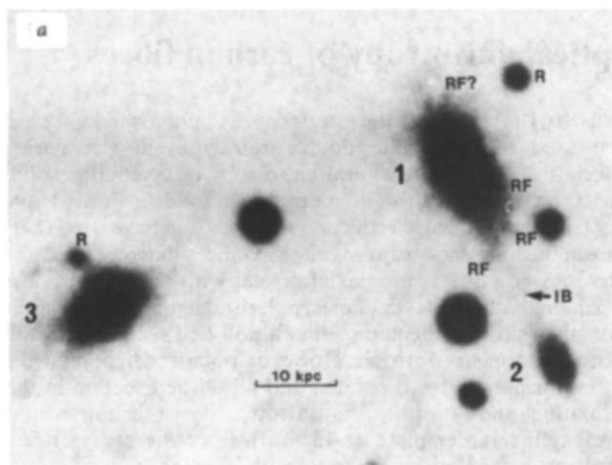


Fig. 2 Enlargements of the peculiar system from the red (a) and blue (b) prints of the Palomar Sky Survey; on a different copy of the survey galaxies 1 and 2 appeared interconnected by a straight red bridge—see text: IB, position of interconnecting bridge; RF, red filaments; R, particularly red stars; B, particularly blue stars. The scale is based on $H = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$. (Reproduced with permission from the National Geographic Palomar Observatory Sky Survey).

A measurement of the Rydberg constant

THE Rydberg constant is of key importance in spectroscopy and, because it may be expressed as the product of several fundamental physical constants, it also plays an important part as an auxiliary constant in the evaluation of the fundamental constants of physics¹. We report here a value of the Rydberg constant that is derived from the measurement of three of the fine-structure components of the hydrogen Balmer- α line by the saturable absorption technique². The earlier measurements of the Rydberg constant by conventional spectroscopy and by the laser saturation absorption spectroscopy technique have been reviewed by Series^{3,4}.

A c.w. tunable dye-laser (CR-599), pumped by the 514 nm line of an argon-ion laser (SP-170), was used as the source of monochromatic radiation. The dye-laser beam was split into counter-propagating saturating ($\sim 10 \text{ mW}$) and probe beams ($\sim 4\%$ of the saturating power), about 2 mm in diameter, these crossed at a small angle, about 1 mrad, at the centre of a Wood's discharge tube which contained hydrogen gas and residual water vapour. The central portion of the tube was 20 mm in diameter and 250 mm long. The voltage gradient, for a discharge current of 30 mA, at a hydrogen pressure of about 6 Pa (0.05 torr) was 0.5 V mm^{-1} . The dye-laser was adjusted to operate at the Balmer- α wavelength by making rapid measurements with a digital wavemeter⁵. The wavelength of the radiation was measured with a plane Fabry-Perot interferometer⁶, whose plates were silvered to give a finesse of about 35. The length of the etalon was about 97.2 mm and was stabilised to be an integral number of half-wavelengths of the radiation from a helium-neon laser which was locked to one of the fine-structure components of ^{127}I . The wavelengths of these components are accurately known in terms of the ^{86}Kr definition of the metre.

Table 1 Values obtained for the reciprocal wavelengths of the fine-structure components and the random, systematic and total uncertainties in the last digit quoted

Fine-structure component	Wavenumber (m ⁻¹)	Uncertainties (s.d.)			No. of observations
		Random	Systematic	Total	
2P _{3/2} -3D _{5/2}	1,523,307.0507	30	104	108	39
2S _{1/2} -3P _{3/2}	1,523,336.5129	61*	135	148	13
2P _{1/2} -3D _{3/2}	1,523,340.0281	65*	130	145	13

* These random errors are correlated

Signals from a confocal Fabry-Perot interferometer (125 MHz fsr) and the plane Fabry-Perot interferometer, together with the saturated absorption signal, were digitised and recorded. A computer was used to fit fringe intensity curves to the etalon signals, and also to fit combined lorentzian and gaussian distributions to the saturated absorption signal. Corrections were made to the measured wavelengths in order to allow for the phase-shift at the etalon mirrors⁶ (-2.6 p.p.b., that is -2.6 parts in 10⁹), Stark shifts⁷ (4.6 p.p.b., -3.3 p.p.b., and 1.3 p.p.b.), a correction for a possible pressure shift (4.4 p.p.b.) and for the unresolved cross-over signals⁸ (-4.2 p.p.b., -8.4 p.p.b. and -13.8 p.p.b.), where the corrections for the three components have (where different) been given in the same order as in Table 1. The results were investigated for possible systematic dependence on the discharge conditions and laser power and, although no significant effects were found, an allowance was made for possible dependence on these parameters in making the systematic error assignments. The width and the asymmetry (which averaged to zero) of the resonances were also studied⁹.

Table 2 Values obtained for the Rydberg constant since 1969

	Rydberg constant less 10,973,730 m ⁻¹	Standard deviation uncertainty (10 ⁻³ m ⁻¹)
Taylor <i>et al.</i> (1969) (review value)	1.20	1,100
Masui (1971)	1.88	800
Kessler (1973)	2.08	850
Kibble <i>et al.</i> (1973)	2.60	800
Hansch <i>et al.</i> (1974)	1.43	100
Goldsmith <i>et al.</i> (1978)	1.476	32
Present work	1.513	85

The values obtained for the wavenumbers of the three fine-structure components studied are given in Table 1, together with the random and systematic uncertainties (expressed as standard deviations) of the last digit quoted. These three values were ratioed with those calculated by Erickson (ref. 10 and personal communication) and then combined with his assumed Rydberg value to give a weighted mean value of 10,973,731.513(85) m⁻¹ for the Rydberg constant; our three Rydberg values agreed to 0.003 p.p.m. The standard deviation uncertainty assigned to our Rydberg measurement consists of the random uncertainty of the mean of 0.0021 m⁻¹ (55 independent results) which has been combined with the root sum square of the systematic uncertainties, expressed as a standard deviation, of 0.0082 m⁻¹. Our result is in excellent agreement with other measurements of the Rydberg constant that have been made during the past 10 years^{8,11-15}, (see Table 2 and ref. 4), particularly the saturable absorption measurement of Hansch *et al.*⁸ and the laser polarisation spectroscopy measurement of Goldsmith *et al.*¹⁵.

We thank R. E. Shawyer for assistance with the computer analysis of the data, and W. R. C. Rowley and A. J. Wallard for

helpful discussions and the loan of an iodine stabilised helium-neon laser.

B. W. PETLEY
K. MORRIS

*Division of Quantum Metrology,
National Physical Laboratory,
Teddington, Middlesex, UK*

Received 15 February; accepted 22 March 1979.

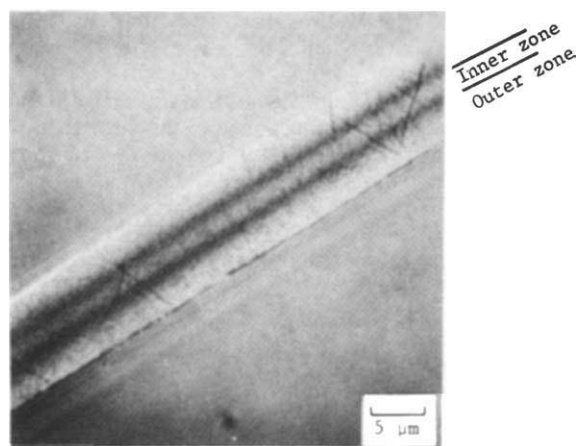
1. Cohen, E. R. & Taylor, B. N. *J. phys. Chem. Ref. Data* **2**, 663-734 (1973).
2. Hansch, T. W., Shahin, I. S. & Schawlow, A. L. *Nature phys. Sci.* **235**, 63-65 (1972).
3. Series, G. W. *Precision Measurement and Fundamental Constants* (ed. Langenberg, D. N. & Taylor, B. N.) 73-82 (NBS Spec. Publ. 343, Washington, D.C. 1971).
4. Series, G. W. *Contemp. Phys.* **15**, 49-68 (1974).
5. Petley, B. W. & Morris, K. *Opt. Quant. Electron.* **10**, 277-278 (1978).
6. Blaney, T. G. *et al. Proc. R. Soc. A* **355**, 89-114 (1977).
7. Blackman, J. A. & Series, G. W. *J. Phys. B* **6**, 1096-1099 (1973).
8. Hansch, T. W., Nayfeh, M. H., Lee, S. A., Cutry, S. M. & Shahin, I. S. *Phys. Rev. Lett.* **32**, 1336-1340 (1974).
9. Petley, B. W., Morris, K. & Shawyer, R. E. (to be published).
10. Erickson, G. W. *J. Phys. Chem. Ref. Data* **6**, 831-869 (1977).
11. Taylor, B. N., Parker, W. H. & Langenberg, D. N. *Rev. Mod. Phys.* **41**, 375-496 (1969).
12. Masui, T. *Natn. Bur. Sids Publ.* **343**, 83-85 (1971).
13. Kessler, E. G. *Phys. Rev.* **7A**, 408-414 (1973).
14. Kibble B. P., Rowley, W. R. C., Shawyer, R. E. & Series, G. W. *J. Phys. B* **6**, 1079-1088 (1973).
15. Goldsmith, J. E. M., Weber, E. W. & Hansch, T. W. *Phys. Rev. Lett.* **41**, 1525-1528 (1978).

Optical anisotropy of carbon fibres

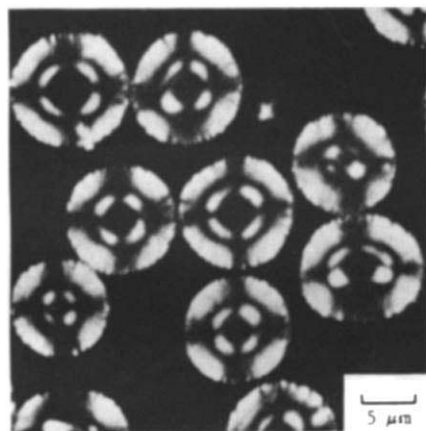
TECHNIQUES based on the interference of polarised light have been used to investigate double refraction in transparent materials due to the structural anisotropy of crystalline solids and strain birefringence in amorphous solids¹. Forrest and Marsh have discussed^{2,3} reflection methods and have used them to examine the anisotropy of the opaque solids carbons and cokes. As graphite is a uniaxial crystal with an optical *c*-axis perpendicular to the layer planes, polarised light incident on the layer-planes is not affected so that a polished section with this orientation appears isotropic. However, polarised light reflected by the prismatic faces is rotated and a polished section in this orientation shows optical anisotropy. By inserting a one wavelength retarder plate at 45° in front of the analyser of a polarising reflection-microscope with crossed polariser and analyser, interference colours are produced from which the orientation of the surface crystallites can be determined.

Polished cross-sections of carbon fibres made from meso-phase pitch and examined in polarised light have a narrow 'Maltese-Cross' pattern with interference colours, deep yellow and blue, in alternate quadrants. By using pyrolytic graphite as a reference material, it can be inferred from the colours that the layer planes are arranged radially in the cross-section. Thin cross-sections which were cut and examined in a transmission electron microscope have confirmed this structure.

High modulus carbon fibres made by the RAE process⁴ from textile PAN fibres (a polyacrylonitrile copolymer) also show a 'Maltese-Cross' but the pattern and the interference colours depend on the diameter of the PAN precursor fibre and the conditions used to convert it to a carbon fibre. The PAN fibres are first oxidised⁴ under tension at temperatures in the range 200-300 °C. At the lower temperature, about 200 °C, the rate of chemical reaction is slow and the rate of diffusion is sufficiently fast to prevent a significant oxygen gradient. At higher temperatures, ≥220 °C, the reaction rate increases so that the oxidation becomes diffusion controlled and if oxidation is incomplete there is^{5,6} a core with little oxidation surrounded by a more oxidised sheath. Longitudinal section (Fig. 1a) of a carbon fibre made after incomplete oxidation shows two zones and cross-sections (Fig. 1b) show two concentric interference figures separated by a dark (isotropic) ring, and having a centre which is also isotropic. From the interference colours these patterns have usually been interpreted⁷ as an inner zone of radially orientated



a



b

Fig. 1 Polished sections of carbon fibres made from 3 denier polyacrylonitrile copolymer fibre by heating in air for 6 h at 220 °C, then carbonising and heating to 2,500 °C. Polariser and analyser crossed. *a*, Longitudinal section; *b*, cross-section.

layer-planes surrounded by an outer zone of layer-planes with circumferential orientation. Extensive oxidative etching in a plasma is said⁸ to support this interpretation.

But no mechanism for the formation of the radial-circumferential structure has been proposed; and investigations using electron microscopy have provided positive evidence that layer-plane orientations in cross-sections of carbon fibres are random. Thus longitudinal sections cut from carbon fibres, made as described above and which in the optical microscope appeared to have a sheath and core, have always appeared uniform in the electron microscope; although selected area diffraction and (002) dark field images (Fig. 2*a*) showed that layer planes of the crystallites were parallel to the fibre axis, there was no preferred distribution between the two zones⁹. Moreover, tilting the specimen produced no overall change; radial-circumferential structures would have had pronounced effects. Transverse sections, ~50 nm thick, have recently confirmed that there was no preferred orientation of the layer planes in the fibre cross-section. Figure 2*b* shows that there were no obvious differences between the phase-contrast lattice-fringe images obtained from the core and sheath regions, and the continuous rings of the selected area diffraction pattern show that the orientations were random.

This raises the possibility that the optical anisotropy seen in the carbon fibres was caused not by structural anisotropy but by strain birefringence, and several recent observations support this.

First, optical activity and interference colours have been seen in reflected polarised light from the glassy carbon matrix in which carbon fibres had been embedded (see Fig. 3*a*); the glassy carbon was made from furfuryl alcohol resin carbonised to 1,000 °C only, a temperature at which stress graphitisation could not have occurred¹⁰. (Similar optical activity has been observed at the stressed interfaces formed when pieces of furfuryl alcohol resin-carbonised to 450 °C were coated with further resin and then carbonised at 1,000 °C.)

Second, cutting thin cross-sections of the carbon fibres caused knife damage; the shattered appearance (Fig. 3*b*) suggests differential residual stresses in the outer and inner zones.

Third, etching with argon ions for short erosion times, <60 min, caused similar amounts of erosion in the sheath and core and no features were seen to indicate any differences in structure even at magnifications of ×20,000. Figure 4*a, b* shows both transverse and longitudinal sections, in which the lower erosion rate of the material between the sheath and core may have been due to its being unstressed.

It is, therefore, suggested that the patterns in both the glassy carbon matrix and in the carbon fibres themselves are caused by strain birefringence. Furfuryl alcohol resin is thermosetting and shrinks by about 20% when heated to 1,000 °C; it forms a rigid cross-linked structure with no plastic flow¹¹ when it is carbonised. This shrinkage will cause a circumferential tensile stress in the matrix around a fibre. Also, it is thought that a radial oxygen gradient in incompletely oxidised PAN fibres may induce internal strains in the resulting carbon fibres. The carbon yield¹² from

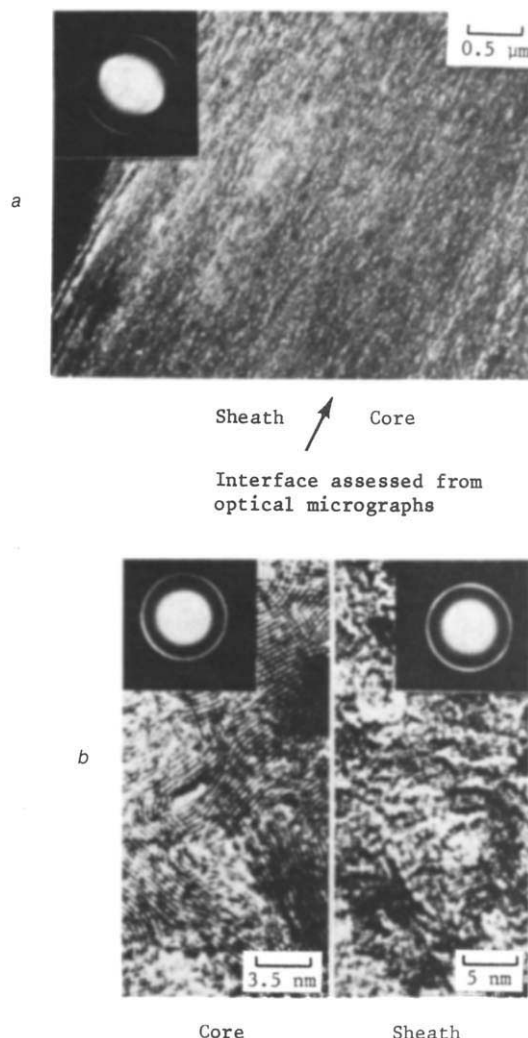


Fig. 2 *a*, (002) Dark field electron micrograph of a thin longitudinal section. *b*, Electron diffraction and lattice-fringe images of a thin cross-section.

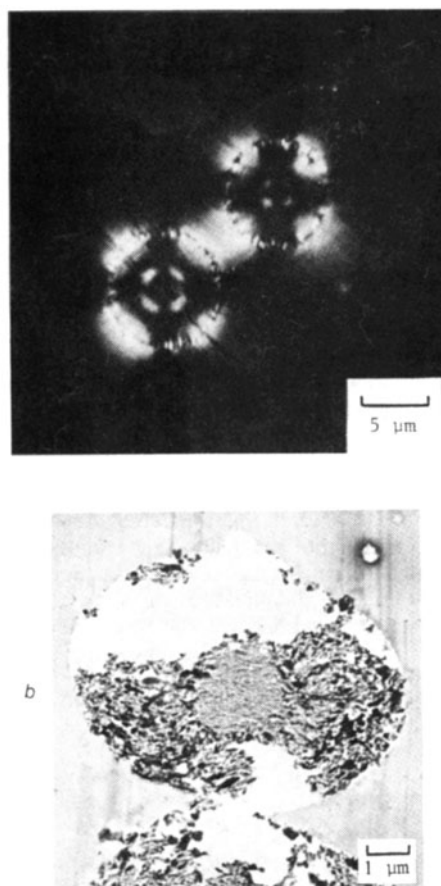


Fig. 3 *a*, Polished cross-sections of carbon fibres in glassy carbon. Polariser and analyser crossed. *b*, A thin cross-section of a carbon fibre.

oxidised PAN fibres is about 50%, and that from unoxidised fibres about 30%; thus any differences in oxygen content will lead to density differences and probably also to differences in the transverse Young's modulus (the dependence of longitudinal Young's modulus on oxygen content is well established)⁴. When a partially oxidised PAN fibre is carbonised there will be a greater carbon yield in the oxygen-rich sheath than in the core. As the fibres do not melt or flow, as they are carbonised, the core will be restrained from shrinking by the more densely packed sheath so that a radial tensile stress will develop in the core with a corresponding circumferential compressive stress in the sheath. The region between the sheath and core is not stressed and seems optically isotropic.

The strain field of a carbon fibre embedded in a glassy carbon matrix (as in Fig. 3*a*) has cylindrical symmetry; hydrostatic stresses around and within each fibre lead to the shear strains shown in Fig. 4*c*.

It is suggested that the presence of a 'Maltese-Cross' pattern in the core of a fibre will depend on a non-uniformity of Young's modulus of the core carbon. If the Young's modulus were constant, the hydrostatic tensile stresses would deform the core evenly; each square of a grid in the core of Fig. 4*c* would be enlarged without distortion and the core would appear optically isotropic in the polarising microscope. But, if there had been an oxygen gradient then the transverse Young's modulus of the core carbon would decrease towards the centre of the fibre, the deformation of the core would be non-uniform and the maximum elongation would be circumferential giving the observed optical anisotropy, the pattern of the core having the same interference colours as that of the glassy carbon matrix; the distortion towards the centre of each fibre is too small to have an effect and it appears optically isotropic. This model also explains why the 'Maltese-Cross' pattern is less evident in micrographs of

carbon fibres made from well-oxidised PAN fibres as it is not seen when the oxidation kinetics do not lead to a significant oxygen gradient.

Transmission electron microscopy of thin transverse sections of carbon fibres made from PAN has shown that the structure is random; it is suggested that the observed optical anisotropy is due to stress in the fibres caused by differential shrinkage between the sheath and core. The optical anisotropy of carbon fibres made from different PAN precursor fibres often differs from a 'Maltese-Cross' and thin-sections of these are being examined.

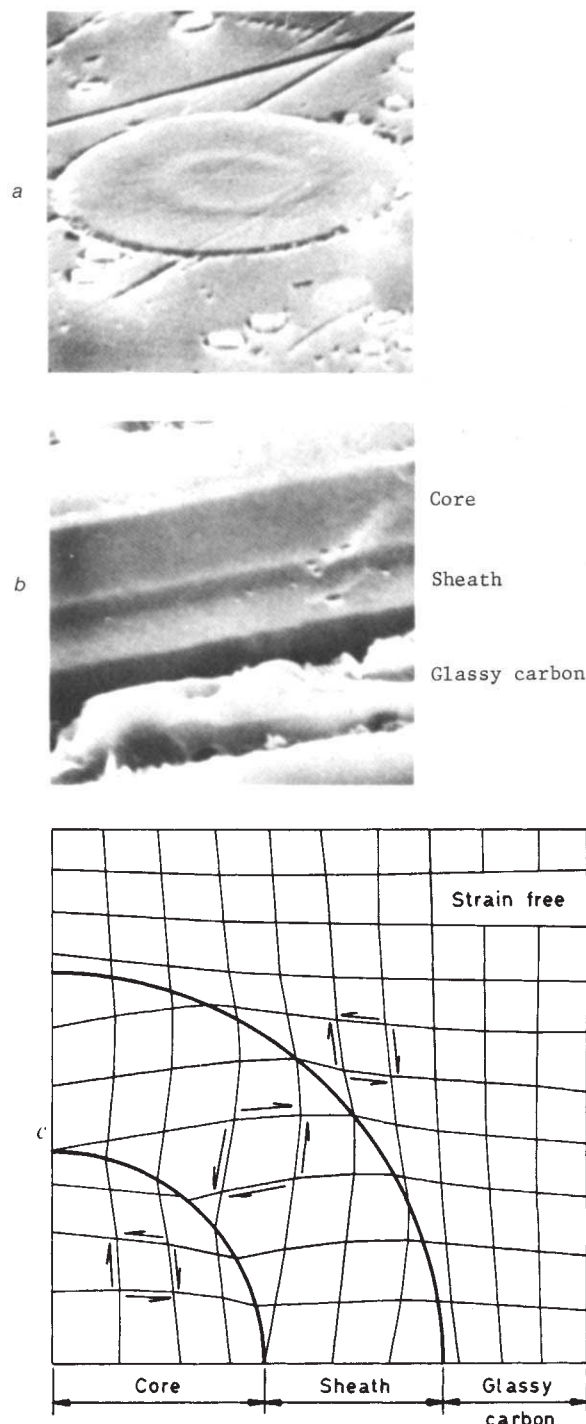


Fig. 4 Polished sections of carbon fibres in glassy carbon argon ion etched; *a*, cross-section fibre diameter 11.8 µm, etched 30 min; *b*, longitudinal section fibre diameter 10 µm, etched 60 min; *c*, a quadrant of the distortion in and around a carbon fibre embedded in a glassy carbon matrix.

I thank Dr S. C. Bennett and Dr D. J. Johnson for the electron-microscopy of thin sections of carbon fibres from pitch and PAN.

W. JOHNSON

Materials Department,
Royal Aircraft Establishment,
Farnborough, Hampshire, UK

Received 11 December 1978; accepted 3 April 1979.

1. Javornickij, J. *Photoplasticity* (Elsevier, Amsterdam, 1974).
2. Forrest, R. A. & Marsh, H. *Carbon* **15**, 349–350 (1977).
3. Forrest, R. A. & Marsh, H. *Proc. 5th Carbon & Graphite Conf.* **1**, 328–333 (1978).
4. Watt, W. & Johnson, W. *Proc. 3rd Carbon & Graphite Conf.* (1970).
5. Watt, W. & Johnson, W. *Nature* **257**, 210–212 (1975).
6. Love, G., Cox, M. & Scott, V. D. *Mater. Res. Bull.* **10**, 8 (1975).
7. Knibbs, R. H. J. *J. Microsc.* **94**, 273–281 (1971).
8. Barnett, F. R. & Norr, M. K. *Carbon* **11**, 281–288 (1973).
9. Bennett, S. C. & Johnson, D. J. *Carbon* (in the press).
10. Kamiya, K. & Inagaki, M. *Carbon* **9**, 287–289 (1971).
11. Kipling, J. J. & Shooter, P. V. *Carbon* **4**, 1–4 (1966).
12. Watt, W. & Johnson, W. *CASI Trans.* **2**, 81–84 (1969).

Evidence for a widespread late Pleistocene humid period in the Kalahari

THE nature and extent of late Quaternary environmental changes in the interior of southern Africa are obscure, in part because of a lack of absolutely dated sediments, and landforms of clear palaeoenvironmental significance. Here I present evidence to show that, during the period 17,000–15,000 BP, the Kalahari, now a semi-arid area with little or no surface drainage, was a region of widely distributed small lakes, as a result of a substantial increase in rainfall.

A belt of pans (small dry or ephemeral lakes) with areas between 0.3 and 7 km² extends along the main Kalahari watershed, between latitudes 23–26°S (Fig. 1). The pans are situated

Table 1 ¹⁴C dates for Urwi pan stromatolites

Pta 2212	15,790 ± 110 BP
Pta 2213	15,610 ± 125 BP
Pta 2146	16,000 ± 160 BP
Pta 2150	16,255 ± 700 BP

in small, shallow, isolated enclosed depressions floored by calcareous and saline clays and sandy clays, laid down in shallow lacustrine conditions. Crescentic dunes on the southern margins of the pans indicate a deflation origin for the depressions¹. The pans thus represent a microcosm of late Quaternary environmental change in the region, with deflation in arid or sub-arid conditions alternating with humid lacustrine conditions.

¹⁴C dates (Table 1) obtained from stromatolites recovered from a site 1–1.5 m above the present level of Urwi pan (Fig. 2) indicate that shallow lacustrine conditions prevailed in the pans 17,000–15,000 yr ago. The stromatolites are of spheroidal morphology, indicating their formation in permanent, wave agitated waters². Changes in their internal structure suggest increased wave action during the period of their formation. The stratigraphic position of the stromatolites (Fig. 2) indicates that they pre-date the last phase of aeolian activity at the pans, and probably represent the final stages of full lacustrine conditions. Complementary sedimentary evidence¹ suggests that permanent alkaline lakes 2–3 times the present area of the pans were succeeded by increasingly arid conditions, in which saline clays were deposited in rapidly contracting seasonal or ephemeral lakes, on the margins of which a sandy wash was laid down. Subsequent deflation of the sandy deposits resulted in the formation of the inner dunes.

Today, mean annual rainfall in the area totals 300–400 mm, with pan evaporation exceeding 3,000 mm. It is thus not surprising that water bodies are ephemeral and runoff from the permeable surface sands negligible (<2 mm per yr). The existence of permanent lakes covering areas 2–3 times the area of

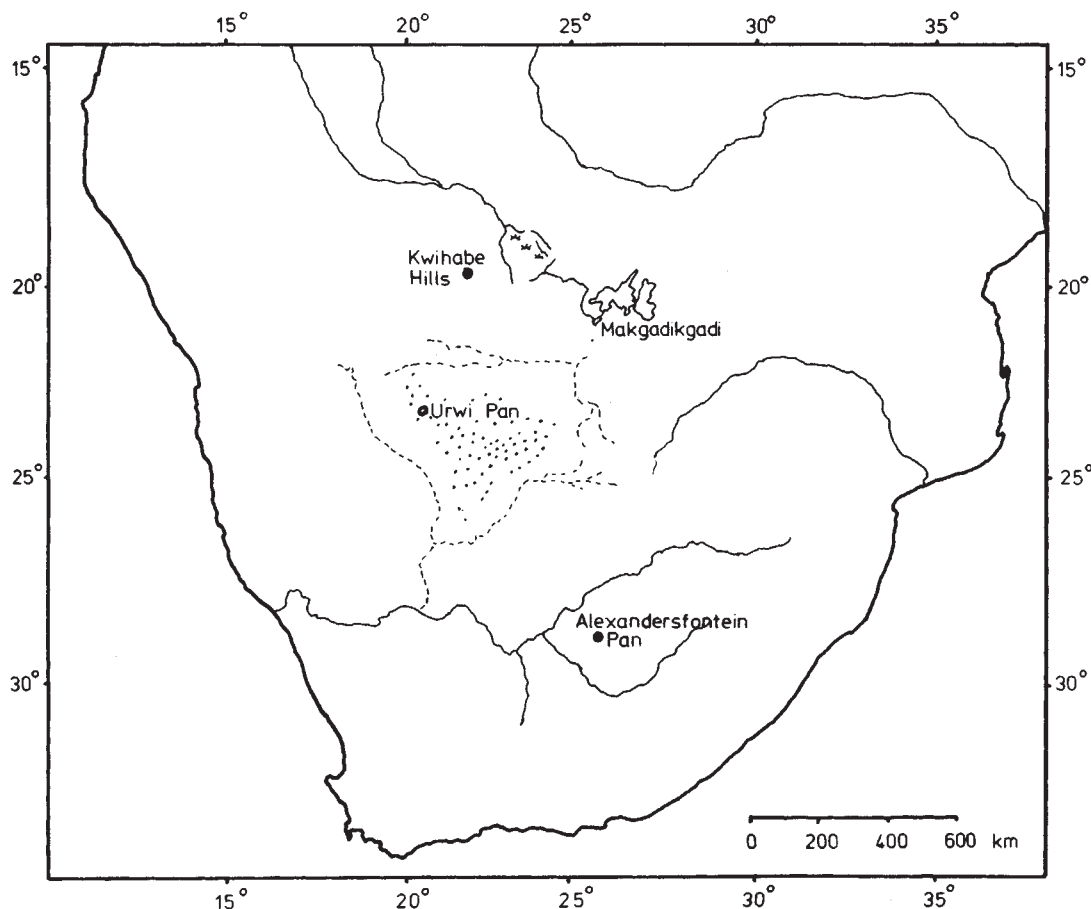


Fig. 1 Southern Africa, showing location of pans, and sites mentioned in text.

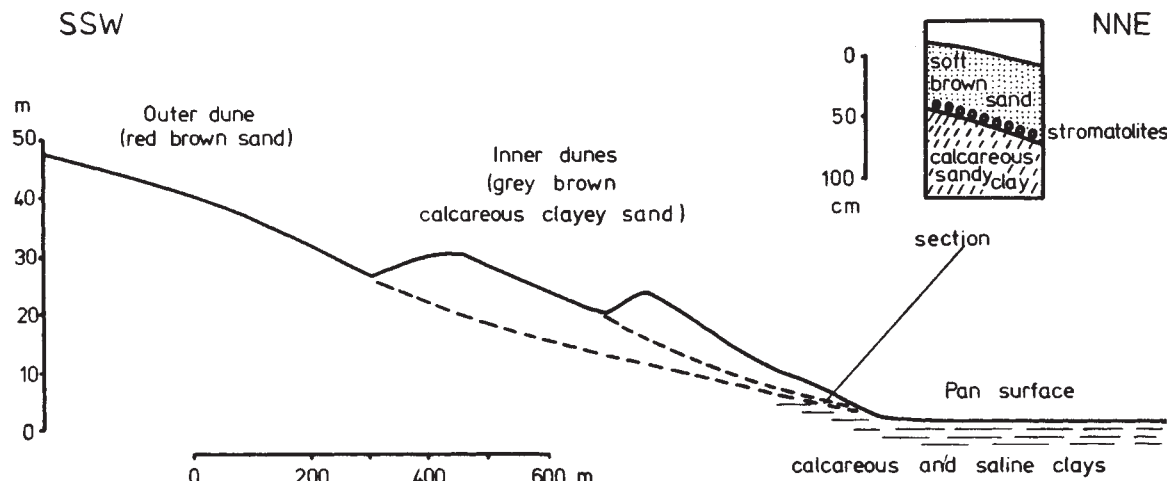


Fig. 2 Stratigraphic section at Urwi pan with location of stromatolites. All pans in region have an essentially similar stratigraphy.

the present pans is clear evidence that the Kalahari experienced substantially increased late Glacial rainfall, reduced evaporation, or a combination of both. Other workers in southern Africa³ have assumed a 5–6 °C lowering of temperature at this time. By analogy with other stations in the interior of southern Africa with a mean annual temperature of 15–16 °C, compared to the 20–21 °C experienced in the southern Kalahari today, evaporation rates 17,000–15,000 yr ago would have been 40–50% lower. Similar comparisons, using data in Midgeley and Pitman⁴ suggest that decreases in temperature and evaporation of this order would have led to increases in surface runoff to 10–20 mm per yr. Even allowing for this, rainfall increases of 1.5–2 times present amounts would have been necessary to sustain permanent lakes in the region, as the catchment areas for the lakes are extremely small (at most 10–15 km²). The potential contribution from shallow groundwater seepage is difficult to assess, but Verhagen *et al.*⁵ demonstrate that significant recharge to shallow groundwater is taking place today in the northern Kalahari, where rainfall totals are 450–600 mm per yr. The increases in rainfall suggested must therefore be regarded as a maximum.

The evidence from the pans thus considerably extends the area of southern Africa known to have experienced increased late Glacial rainfall, and strongly suggests that the climate of the whole of the interior of the subcontinent was sub-humid at this time. Then the arid zone in southern Africa must have largely disappeared, except for isolated refuges along the Namib coast.

Butzer *et al.*³ have calculated that rainfall totals were 123% of present amounts 16,000 yr ago in Alexandersfontein pan area near Kimberly (28°50' S); Cooke⁶ estimated a threefold increase in rainfall for the Kwihahe Hills area in northwestern Botswana (20° S), based on cave deposits with ¹⁴C ages ranging from 14,000 to 17,000 BP. High lake levels in the Makgadikgadi depression some 20,000 years BP⁷ may indicate that the period of increased rainfall started some 5,000 yr before this.

The picture of late Glacial climatic conditions in southern Africa which is now emerging is significantly different to that of the rest of the continent, which seems to have been dry at this time, with the southern margins of the Sahara up to 500 km south of their present position and low lake levels in east Africa; only parts of north Africa experienced wetter climates during this period.

The increased rainfall in the Kalahari at this time was probably the result, not of increased penetration of the subcontinent by winter cyclonic rains as Butzer *et al.*³ suggest, but of greater summer rainfall due to a more southerly position of the Inter Tropical Convergence Zone. This was in turn the result of strengthened Northern Hemisphere circulations. Modern

parallels support this argument, with heavy rainfall in the region in recent years (1972, 1974–5, 1977–8) being associated with the persistence of the Inter Tropical Convergence Zone near its southerly limits (19–20° S). Desiccation of the interior of southern Africa was not well advanced until 9,000 BP, and was associated with much higher temperatures⁸, at a time when rainfall over much of Africa between the Equator and the northern Tropic was substantially greater than at present⁵. Equally interesting is the lack of correlation with Australia, where the late Glacial was cold, dry and windy⁹.

I thank Dr J. C. Vogel for ¹⁴C determinations on the stromatolites.

I. N. LANCASTER

Department of Geography and Environmental Studies,
University of the Witwatersrand,
Johannesburg, South Africa

Received 23 November 1978; accepted 13 March 1979.

1. Lancaster, I. N. *Geogr. J.* **144**, 81–98 (1978).
2. Logan, B. W., Rezak, R. & Ginsburg, R. N. *J. Geol.* **72**, 68–83 (1964).
3. Butzer, K. W., Fock, G. J., Stuckenrath, R. & Zilch, A. *Nature* **243**, 328–330 (1973).
4. Midgeley, D. C. & Pitman, W. V. *Hydrological Research Unit, University of the Witwatersrand Report 2/69* (1969).
5. Verhagen, B. T., Mazor, E. & Sellschop, J. P. F. *Nature* **249**, 643–644 (1974).
6. Cooke, H. J. *Geogr. J.* **141**, 177–202 (1975).
7. Street, F. A. & Grove, A. T. *Nature* **261**, 385–390 (1974).
8. Coetzee, J. A. *Paleoecol. Afr.* **3**, 100–123 (1967).
9. Roggon, P. & Williams, M. A. J. *Paleogeogr. Paleoclim. Paleocool.* **21**, 285–327 (1977).

Tracing the wake of a flying bird

STUDY of the classical problem concerning “the way of an eagle in the air”, when applied to flapping bird flight, offers many difficulties. All dynamic interactions between an object and a fluid medium in which it moves are ciphered in the structure of its wake, but the wake behind a flying bird is difficult to work with because it is an invisible and very short-lived formation; moreover, its structure must be complicated as the bird's wings are working in a constantly changing manner. I report here the results of wake visualisation experiments carried out during short flights of a chaffinch (*Fringilla coelebs*) and a brambling (*F. montifringilla*) in enclosures. The general configuration of a wake is given, but no measurements have been made: in general, to extract some quantitative information from the wake structure, a more precisely documented picture of the dynamics of its formation is needed.

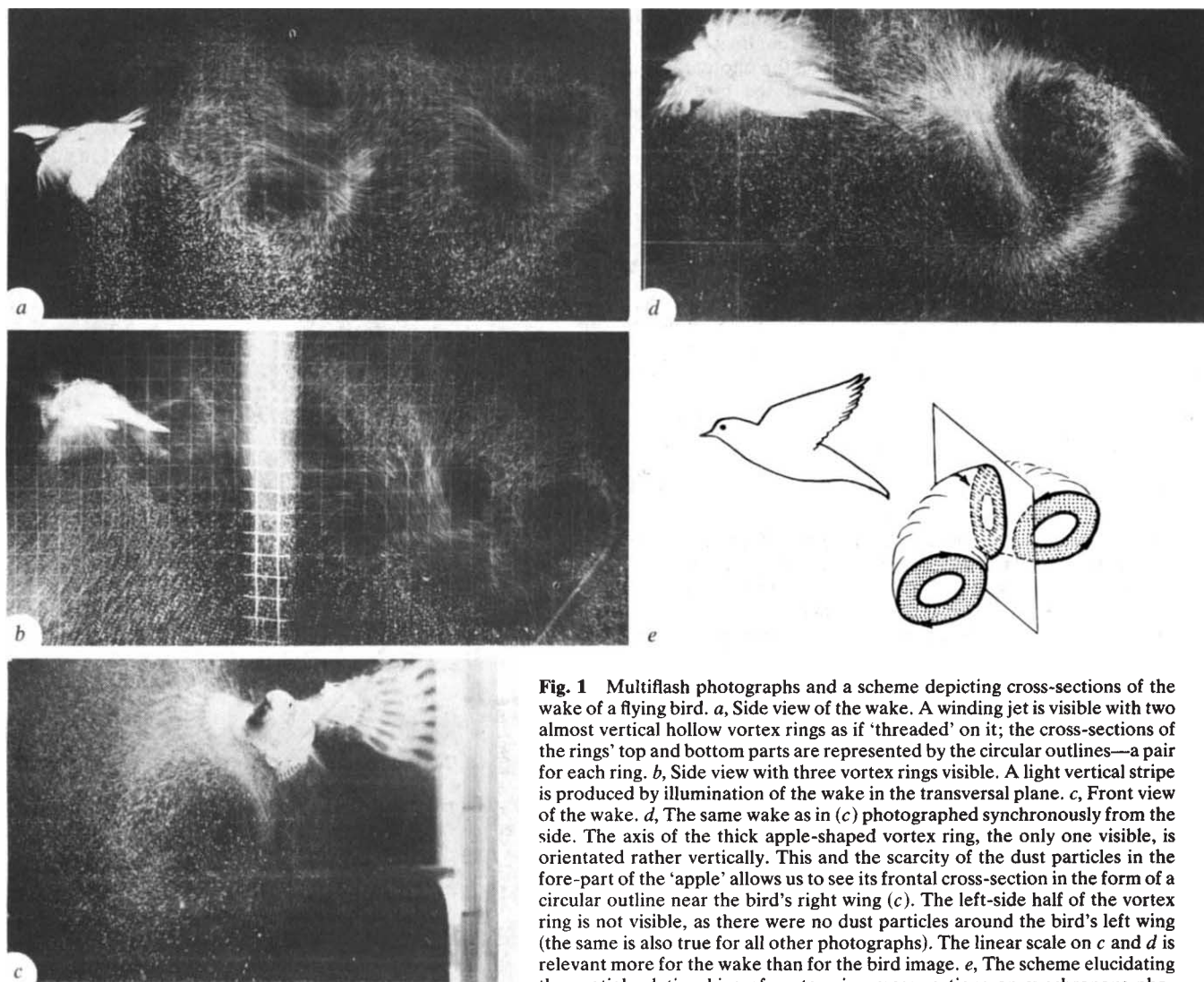


Fig. 1 Multiflash photographs and a scheme depicting cross-sections of the wake of a flying bird. *a*, Side view of the wake. A winding jet is visible with two almost vertical hollow vortex rings as if 'threaded' on it; the cross-sections of the rings' top and bottom parts are represented by the circular outlines—a pair for each ring. *b*, Side view with three vortex rings visible. A light vertical stripe is produced by illumination of the wake in the transversal plane. *c*, Front view of the wake. *d*, The same wake as in (*c*) photographed synchronously from the side. The axis of the thick apple-shaped vortex ring, the only one visible, is orientated rather vertically. This and the scarcity of the dust particles in the fore-part of the 'apple' allows us to see its frontal cross-section in the form of a circular outline near the bird's right wing (*c*). The left-side half of the vortex ring is not visible, as there were no dust particles around the bird's left wing (the same is also true for all other photographs). The linear scale on *c* and *d* is relevant more for the wake than for the bird image. *e*, The scheme elucidating the spatial relationships of vortex ring cross-sections on synchronous photographs (*c, d*); as in the photographs only the right-side half of the vortex ring is shown. The cross-sections in the plane of the drawing correspond to the side view of the wake (*d*) and the one at right angles to this plane represents the front view of the wake (*c*). The direction of the vortex ring circulation is indicated by the arrows. The pictures were taken in a net cage (*a, b*) and in a large aquarium (*c, d*); the birds photographed were brambling (*b*) and chaffinch (*a, c, d*); the number of flashes per series was about 13 (*a*), 8 (*b*) and 7 (*c, d*); flash frequencies were 1 kHz (*a, b*) and about 0.75 kHz (*c, d*); visualisation was carried out with paper (*a, b*) and wood (*c, d*) dust.

The method involved multiple flash photography of moving small light particles forming a cloud through which the bird was forced to fly. Several visualising agents were tried but wood and paper dusts manufactured on a grindstone were most successful. Mean diameters of individual particles of both types of dust were 0.5–0.6 mm (paper dust had a more pronounced tendency to aggregate in flakes of a mean diameter of ~ 1.0 mm), but average sinking velocities were different: 0.5 m s^{-1} in wood and 0.2 m s^{-1} in paper dust. Birds performed their short flights between two perches in a Plexiglass aquarium ($2 \times 0.5 \times 1 \text{ m}$) and in a net cage ($2.6 \times 0.6 \times 0.7 \text{ m}$). An electronic generator producing a series of flashes was designed for the study. The generator was used in a waiting regime and was triggered from one or two photorelays as the bird crossed an infrared beam. The frequency of flashes and their number per series were widely adjustable. Satisfactory results were obtained with frequencies of the order of 750–1,000 Hz and 6–8 flashes per series.

The experimental procedure was as follows. The room was darkened (only very faint illumination being retained) and the shutters of two photographic cameras were opened. Then a cloud of flow-visualisation particles was blown out with the aid of a long rubber tube from a container attached above the middle part of a bird's trajectory. At the same time the bird was flushed from the perch. The flying bird triggered the flash

generator and hence a wake photograph was produced. Cameras were then prepared for the next exposure and the cycle repeated.

In the photographs obtained the multiple images of individual particles clearly show directions and relative velocities of their movement in the plane of the paper. The faster a particle is moving the longer the segments of a broken line it produces. One can consider these lines as an approximation to the real trajectories of fluid particles. With the linear scale and flash frequencies known it is possible to calculate roughly a number of characteristic velocities in the plane of the wake flow field pictured.

To analyse the photographs obtained (Fig. 1) it is necessary to consider some peculiarities of the methods used. Photographic superposition of successive images causes a strong impression as if light objects, specifically dust particles, are always situated in front of dark ones. The bird itself moves relatively slowly, so its image is usually overexposed with few details visible—exposures being chosen to reproduce primarily the wake structure. Under the action of a centrifugal force the core parts of rotating bodies of air are free from visualising material. This causes accentuation of such structures, but it should be remembered that their apparent dimensions are also strongly dependent on the size and density of visualising particles. Finally, the

particle cloud was formed by the experimenter in such a way that it embraced only the wing of the bird remote from the observer, so that sections of the wake were produced on the photographs.

The two-dimensional longitudinal view of the bird wake conforms reasonably well to the reversed vortex street with a 'reactive' jet between the vortices—the pattern that was predicted long ago by theorists and then repeatedly reproduced in experiments with an oscillating aerofoil¹. Analysis of the photographs also reveals that the spatial structure of the wake of an actively flying bird is represented by a series of thick and more or less deformed vortex rings with an intense and winding jet which passes through the rings' holes. Form and orientation (that is, axis inclination) of the vortex rings are variable because the photographs correspond to different stages of birds' short flight with different requirements on lift and thrust and unequal contribution by the ballistic component (the initial jump from the perch) at each stage.

The development of a vortex ring can be conceived in the following way. The upper part of the ring (if it is placed vertically) is formed by the starting vortices of both the wings at the beginning of their downstroke; these are extended on the sides of the ring, during the middle stages of the downstroke, by the wings' tip vortices. The lower part of the ring is formed by the stopping vortices shed from the wings in their lowest position at the bottom of the downstroke. The vortices generated by the left and right wing are closed against each other at the upper and lower points of the wing stroke. So a vortex ring is formed. In most if not all passerine birds all the aerodynamic work is known to be done only on the downstroke; on the upstroke the wings are folded and their interaction with the air is minimal². This is why vortex rings in the wake of the finches studied here are produced only during the downstroke.

A recent theoretical model^{3,4} has predicted the configuration of the vortex rings observed in these experiments as being the most realistic description of the wake of a flying bird, consistent with the unsteady regime of its wing beat. This model provides a framework for calculations of the power consumed in flight and for the quantification of the vortex distribution of the wake.

I thank Professor Sir James Lighthill for the kind attention he has given to this report and V. I. Petrowsky for designing the generator and helping to produce the photographs.

N. V. KOKSHAYSKY

A. N. Severtzov Institute of Evolutionary
Animal Morphology and Ecology,
Leninsky Prospekt 33,
Moscow W-71, USSR

Received 21 February; accepted 5 April 1979.

1. Kokshaysky, N. V. *Ocherk Biologicheskoi Aero- i Gidrodinamiki (Polet i Plavanie Zhivotnykh)* [An Essay on Biological Aero- and Hydrodynamics (Flight and Swimming of Animals)] (Nauka, Moscow, 1974).
2. Brown, R. H. J. *Biol. Rev.* **38**, 460–489 (1963).
3. Rayner, J. M. V. *J. exp. Biol.* **80** (in the press).
4. Rayner, J. M. V. *J. Fluid Mech.* **91**, 739–770 (1979).

Shoot height, weight and standing crop in relation to density of monospecific plant stands

DESPITE great differences in architecture and supporting tissue, monospecific stands of single-stemmed plants ranging in size from a moss to a tree follow a simple rule relating shoot dry weight and standing crop to density. We describe here an equation for this relationship which remains true for almost any fully occupied plant stand dominated by a single species.

Data were obtained, mostly from the literature^{1–19}, for 65 stands of 29 different plant species (Table 1) ranging from mature midsummer shoots of annual and perennial non-woody species to saplings and full-sized trees. Figure 1a shows the

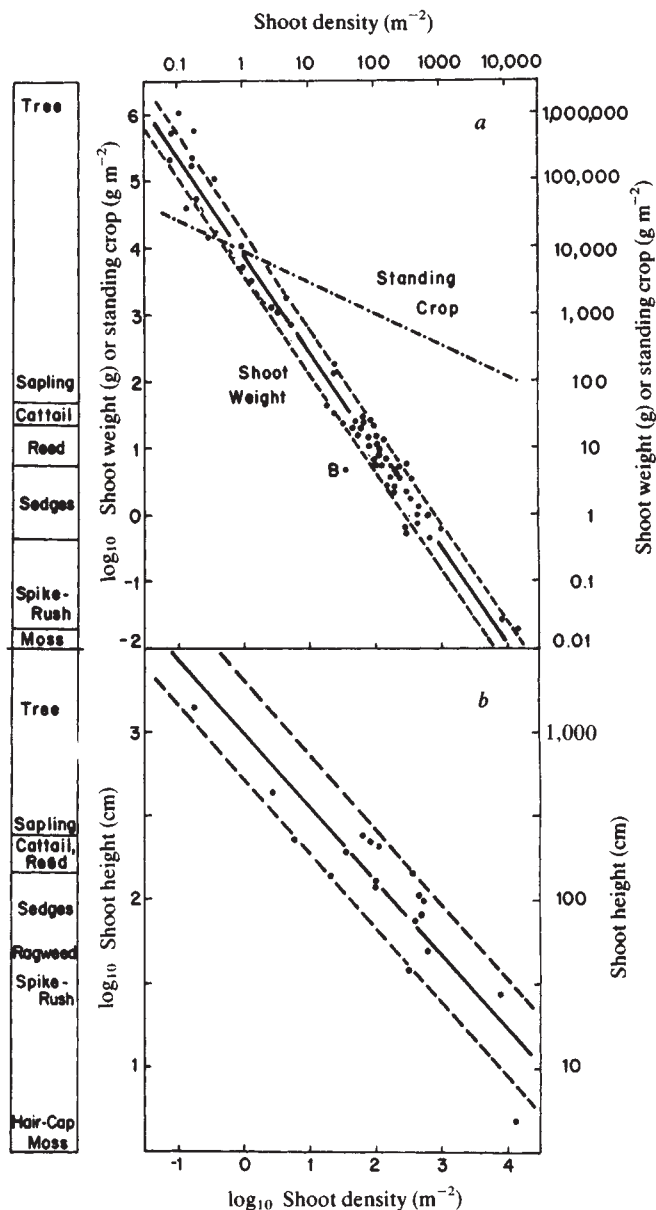


Fig. 1 The relationship of shoot dry weight and standing crop (a) and height (b) to density in monospecific stands of plants. Dashed boundary lines at $0.3 \log_{10}$ intervals below and above the solid lines represent plants half and twice as large. Point B represents *Bolboschoenus maritimus*.

inverse, double logarithmic relationship of shoot dry weight (w , grammes) to density (d , m^{-2}). It is calculated as the reduced major axis^{20,21}, because neither variable is truly dependent on the other, and has the form $w = 9,670d^{-1.49}$ and the correlation coefficient $r_{\log 10} = -0.987$.

A simple geometric model, developed by Yoda *et al.*¹, to explain intraspecific weight/density relationships during self-thinning of overcrowded stands of individual plant species, seems also to apply to interspecific weight/density relationships among diverse species, ranging in shoot weight over more than seven orders of magnitude and in shoot density over five orders of magnitude. The model is described as follows: (1) assume that available space is fully occupied (that is, that the leaf canopy is closed), and that the area occupied by a plant is proportional to the square of some linear dimension (for example, basal diameter or height); (2) then weight (or volume) per shoot will be proportional to the cube of the linear dimension; (3) weight per shoot is therefore proportional to the $3/2$ power of the area occupied or to the $-3/2$ power of its reciprocal, which is density.

This exponent of -1.5 is very close to that observed in Fig. 1a (-1.49), which therefore appears to fit the suggested model despite major differences in plant architecture (for example, *Polytrichum*, *Pteridium*, *Carex*, *Typha*, *Phragmites*, *Ambrosia*, *Betula*, *Abies*), and in the nature of supporting tissues, with a large component of parenchyma in some species (for example, *Eleocharis*, vegetative shoots of *Typha*) and of sclerenchyma in others (*Abies*, *Betula*).

The relationship may be followed further by examining the ranges of shoot height and weight (Table 2). The rules of the simple geometric model dictate that the range for $\log_{10}$ height ought theoretically to be half the range for $\log_{10}$ density, whereas the range for $\log_{10}$ weight ought to be 1.5 times the range for $\log_{10}$ density. The observed ranges are in good agreement with the model, being within 7% of the theoretical range for height, and within 2% of the theoretical range for weight. Conformity would probably be greater if height data were available for the largest *Abies sachalinensis*. It is 6.5 times the weight of *Taxodium distichum*, the largest plant for which height data are available. According to the model, 6.5 times greater weight implies 1.87 times greater height, and an increase of 0.27 in the $\log_{10}$ height range. Adding 0.27 to 2.44 gives 2.71, within 4% of the theoretical height range for all 27 species.

The relationship of shoot height (h , cm) to density (d) of 15 species in 19 stands is shown in Fig. 1b, and is given by the reduced major axis, $h = 970d^{-0.443}$. The correlation coefficient, $r_{\log 10} = -0.902$. Thus, density explains only 81% of the variance in shoot height, as compared with 98% of the variance in weight for the same 19 stands. Marked architectural differences are reflected in the ratios of height to cube-root of dry weight among the different species. Ratios for seven species of *Carex* range from 67 to 130, whereas two species of *Typha* have ratios of 40 and 68. Three stands of *Phragmites communis* range from 76 to 93. Trees exhibit lower ratios, 39 and 26 for *Pinus silvestris* and 26 for *Taxodium distichum*. The moss *Polytrichum commune* has the lowest ratio, at 19.

Table 1 Species investigated*

<i>Abies sachalinensis</i> ¹	<i>Glyceria maxima</i> ²
† <i>Ambrosia artemisiifolia</i> (unpubl.)	† <i>Phragmites communis</i> ^{2,9-13}
<i>Ammophila arenaria</i> (unpubl.)	† <i>Pinus silvestris</i> ¹⁴
<i>Betula</i> (platyphylla, Ermani, Maximowicziana) mixed ¹	† <i>Polytrichum commune</i> (unpubl.)
<i>Bolboschoenus maritimus</i> ²	<i>Pteridium aquilinum</i> ¹⁵
† <i>Carex acuta</i> ⁵	<i>Scirpus fluviatilis</i> ⁸
<i>C. atherodes</i> ⁴	<i>S. validus</i> ⁸
† <i>C. elata</i> ³	† <i>Sparganium erectum</i> ¹⁶
† <i>C. lacustris</i> ^{3,4,5}	<i>S. eurycarpum</i> ⁸
<i>C. lanuginosa</i> ⁴	† <i>Taxodium distichum</i> ¹⁷
† <i>C. lasiocarpa</i> ^{3,4}	<i>Typha angustifolia</i> ^{2,13}
† <i>C. rostrata</i> ^{3,4,6,7,8}	<i>T. glauca</i> ⁸
† <i>C. vesicaria</i> ³	† <i>T. latifolia</i> ^{18,19}
† <i>Eleocharis calva</i> (unpubl.)	† <i>T. (angustifolia, glauca, latifolia)</i> mixed, (unpubl.)

*Other species accounted only for a few % of above-ground standing crop and less than 10% in all measured cases.

†Data for both weight and height were available.

The weight/density relationship within a monospecific plant stand undergoing competitive self-thinning is generally characterised¹ by a negative exponent of -1.5 . Different species are expected^{1,22} to exhibit distinct differences in the constant C of the equation $w = Cd^{-1.5}$. Several species should thus yield a family of curves, all with negative exponents near -1.5 but with different values for the constant C presumably reflecting differences in plant architecture. However, within the limits of variability normally observed for such an inverse power-law relationship, most species fit reasonably well to the solid line in Fig. 1a, defined by a single value (9,670) of the constant C . Threequarters of the 65 plant stands in Fig. 1a lie within the

dashed boundary lines shown 0.3 $\log_{10}$ intervals below and above the fitted solid line. These boundary lines ($C = 4,900$ and $19,500$ respectively) represent shoots one-half and twice as large as those which fit the equation given earlier. Moreover, data for many individual species reveal a variability at least as great as that represented by the dashed boundary lines. Data from laboratory experiments may be similarly variable. Two populations of *Helianthus annuus* allowed to self-thin at full light intensity²² exhibited C values of 6,920 and 91,200 respectively, and regression exponents of -1.33 and -1.84 . The clumped distribution of many plant shoots and the often logarithmic frequency distribution of shoot sizes in overcrowded stands¹ contribute greatly to the substantial variability.

Table 2 The ranges of height*, density† and weight† for the plant shoots in Fig. 1

	Minimum	Maximum	$\log_{10}$ range	
			Actual	Theoretical
Height (cm)	5.0	1.39×10^3	2.44	2.62
Density (m^{-2})	0.079	1.38×10^4	5.24	—
Dry weight (g)	0.019	1.07×10^6	7.75	7.86

*19 stands, 15 species. †65 stands, 27 species.

Architecturally diverse species overlap considerably, and where deviations from the reduced major axis occur they do not seem to be related in any consistent way to differences in plant architecture or in the nature of the supporting tissues. A number of self-thinning curves in the literature^{1,22} also conform rather well to the relationship shown in Fig. 1a, especially if reasonable corrections are made for below-ground material of whole plants. Even if more detailed study reveals systematic variations in the constant C which can be related to differences in plant architecture, Fig. 1a suggests that its values will lie within a fairly narrow range, and mostly between half and twice the value reported here.

Two species, *Chenopodium album*¹ and *Bolboschoenus nigricans* (B in Fig. 1a), appear to deviate substantially from the relationship reported here. The deviation of *Chenopodium* is particularly striking because of the conformity of a similar weed, *Amaranthus retroflexus*¹. However, according to one of the authors (T. Kira), their *Chenopodium* data were for fresh weight, others were for dry weight. *Bolboschoenus* (= *Scirpus*) *maritimus*, marked B in Fig. 1a, lies well below the reduced major axis. This rhizomatous species colonises open water, and probably had not developed a closed canopy. The relationship will also break down where aridity and cold prevent development of a complete plant cover, as in deserts and tundra.

Self-thinning is not always characteristic of monospecific stands of plants in nature, particularly in uncrowded situations. The exponent in the weight/density relationship is generally close to -1.0 in such cases, so that standing crop per unit area is independent of density. This is the 'Law of Constant Final Yield'²³ and this holds for stands of *Taxodium distichum* in Georgia¹⁷ and *Phragmites communis* in Czechoslovakia¹². The reduced major axes for both species lie close, although at an angle, to their respective short segments of the fitted line in Fig. 1a. Apparently this line sets an approximate upper limit to the mean size of plant shoots at any given density, whether or not self-thinning is taking place.

The negative slope in Fig. 1a obviously reflects structural constraints, with a less dense plant spacing allowing a larger basal area for stem support, and hence a greater ultimate height and mass. This structural explanation might be thought to have an underlying basis in the physiological balance between respiration, a function of shoot weight, and photosynthesis, a function of leaf area. In such a case, shoot weight ought to be

closely related to leaf area index. However, leaf area indices calculated from data compiled by R. H. Whittaker²⁴ show little difference in communities with very different shoot densities and sizes. For example, temperate and tropical evergreen and deciduous forests exhibit mean indices between 5 and 12 m² m⁻², as compared to indices of 3.6 for temperate grasslands and 7.0 for swamps and marshes. Grassland and wetland communities have far smaller shoot weights and greater shoot densities than forests. Moreover, fourteen Czechoslovak stands of *Phragmites communis*¹² exhibit no relationship of leaf area index, which ranges from 2.4 to 6.3, to either shoot weight or density, which correlate strongly with one another ($r_{\log 10} = -0.91$). White²⁵ discusses the point in more detail.

It follows from the weight/density relationship in Fig. 1a that large plants growing at low density yield a larger above-ground standing crop than small plants at high density, with $wd = 9,670d^{-0.49}$. This relationship is shown as a dash-dot line in Fig. 1a. It can be seen that standing crop at a density of 0.1 shoots m⁻² (characteristic of a large tree) is 26,000 g m⁻², whereas at a density of 10,000 shoots m⁻² (characteristic of a large moss), it is 120 g m⁻². Only widely spaced plants, which can develop a large basal area and so grow to great heights, are able to accumulate large amounts of biomass. This is so obvious as to be a truism. The remarkable fact is that a single, simple rule relating shoot weight and standing crop to density should apply over so many orders of magnitude and to plants as diverse in their architecture as mosses, ferns, gymnosperms, monocotyledons and dicotyledons.

After submitting this article I learned from J. L. Harper that a similar relationship was described at the Twelfth International Botanical Congress in Leningrad by J. White, who subsequently informed me that it will be dealt with further in a forthcoming publication.

I thank E. C. Abbe, W. L. Koukkari, E. G. and D. J. McLaughlin, P. A. Morrow, D. C. Pratt and A. van der Valk for their comments, and E. J. Cushing for a detailed review. A. van der Valk and W. H. Schlesinger provided data in advance of publication. J. White transmitted T. Kira's explanation of the aberrant *Chenopodium* data, and informed me of his 1977 and forthcoming papers.

EVILLE GORHAM

Department of Ecology and Behavioral Biology,
University of Minnesota,
Minneapolis, Minnesota 55455

Received 20 November 1978; accepted 9 February 1979.

1. Yoda, K., Kira, T., Ogawa, H. & Hozumi, K. *J. Biol. Osaka Cy Univ.* **14**, 107-129 (1963).
2. Ondok, J. P. in *Ecosystem Study on Wetland Biome in Czechoslovakia* (ed. Hejný, S.) 83-85 (Czechosl. IBP/PT-PP Rep. no. 3, Třeboň, 1973).
3. Mörsjö, T. *Opera bot.* No. 24, pp 187 (1969).
4. Gorham, E. & Bernard, J. M. *J. Minn. Acad. Sci.* **41**, 15-17 (1975).
5. Bernard, J. M. & MacDonald, J. G. Jr. *Can. J. Bot.* **52**, 117-123 (1974).
6. Gorham, E. & Somers, M. G. *Can. J. Bot.* **51**, 1097-1108 (1973).
7. Bernard, J. M. *Ecology* **55**, 350-359 (1974).
8. Van der Valk, A. & Davis, C. B. in *Freshwater Wetlands* (eds Good, R. E., Whigham, D. F. & Simpson, R. L.) 21-37 (Academic, New York, 1978).
9. Dykijová, D. & Hradecká, D. *Folia Geobot. phytotax. Praha* **11**, 23-61 (1976).
10. Gorham, E. & Pearsall, W. H. *Oikos* **7**, 206-214 (1956).
11. Haslam, S. J. *Ecol.* **60**, 585-610 (1972).
12. Kvet, J. in *Ecosystem Study on Wetland Biome in Czechoslovakia* (ed. Hejný, S.) 93-95 (Czechosl. IBP/PT-PP Rep. no. 3, Třeboň, 1973).
13. Mason, C. F. & Bryant, R. J. *J. Ecol.* **63**, 71-95 (1975).
14. Ovington, J. D. *Ann. Bot.* **21**, 287-314 (1957).
15. Pearsall, W. H. & Gorham, E. *Oikos* **7**, 193-201 (1956).
16. Dykijová, D. & Ondok, J. P. *Preslia, Praha* **45**, 19-30 (1973).
17. Schlesinger, W. H. *Ecol. Monogr.* **48**, 43-65 (1978).
18. Boyd, C. E. *Bull. Torrey Bot. Club* **98**, 144-150 (1971).
19. McNaughton, S. J. *Ecol. Monogr.* **36**, 297-325 (1966).
20. Imbrie, J. *Bull. Am. Mus. Nat. Hist.* **108**, 211-252 (1956).
21. Till, R. *Statistical Methods for the Earth Scientist* (Wiley, New York, 1974).
22. White, J. & Harper, J. L. *J. Ecol.* **58**, 467-485 (1970).
23. Harper, J. L. & White, J. in *Proc. Adv. Study Inst. Dynamics Numbers Popul.* (eds den Boer, P. J. & Gradwell, G. R.) 41-63 (Oosterbeek, 1970).
24. Whittaker, R. H. *Communities and Ecosystems* 2nd edn (Macmillan, New York, 1975).
25. White, J. *Nature* **268**, 373 (1977).

Parasite pathogenicity and the depression of host population equilibria

THE use of pathogens as agents for the biological control of pest species has increased rapidly in recent years with the realisation that chemical pesticides are not the panacea they were once thought to be¹⁻⁶. The selection of a suitable parasite is often based on the assumption that highly pathogenic organisms will be most effective in minimising pest population growth⁶⁻⁸. Recent theoretical studies of the dynamics of host-parasite associations provide a template for testing this assumption⁹⁻¹². This paper examines the relationship between parasite pathogenicity, as measured by the severity of the pathogen's influence on host survival, and the consequent depression of the host population from the size it would have achieved in the absence of infection. The maximum degree of depression is achieved by direct life cycle parasites of moderate to low pathogenicity, a prediction which has important implications for the use of pathogens as biological control agents. In agreement with traditional beliefs^{13,14}, theory suggests that highly pathogenic species will cause their own extinction but not that of their host. The relationship between population depression and host reproductive potential is shown to be critically dependent on the biological characteristics of the pathogen.

Parasites can be divided into two main classes. Microparasites (viruses, bacteria, protozoans) are characterised by small size, short generation times, extremely high rates of direct reproduction within the host, and a tendency to induce lasting immunity to re-infection in those hosts that survive initial infection¹⁵. The duration of infection is typically short in relation to the expected lifespan of the host, and thus is of a transient nature (there are exceptions to this general trend, as shown, for example, by herpes simplex virus¹⁵ and the slow viruses¹⁶). Macroparasites (parasitic helminths and arthropods) have much longer generation times, and direct multiplication within or on the host is either absent or occurs at a low rate compared with that of microparasites.

Such immune responses as are elicited by these metazoans tend to depend on the number of parasites present in a given host and to be of relatively short duration relative to the expected lifespan of the host¹⁷⁻¹⁹. Macroparasitic infections therefore tend to be of a lasting nature, as the hosts are susceptible to continued re-infection.

Both microparasites and macroparasites may complete their life cycles by passing from one host to the next either directly or indirectly via one or more intermediate host species. Pathogens which have been used in biological control¹⁻³, and those whose potential is now being investigated⁴⁻⁸, all exhibit direct life cycles (various viruses, bacteria, protozoa and nematodes).

The influence of parasite pathogenicity (*per capita* rate of parasite-induced host mortality) on the dynamics of host population growth may be explored by means of population models framed in differential equations. The mathematical theory of the dynamics of microparasitic infectious diseases, as recently reviewed by Bailey²⁰, is largely based on compartmental models which delineate susceptible, infected and immune categories of hosts (containing, respectively, x , y and z numbers of hosts). The model used to examine the dynamics of microparasitic infections differs from standard epidemiological models in assuming that the total number of hosts, $H = x + y + z$, is not treated as some independently set constant, but is assumed to be a variable determined by both the dynamics of the disease and the natural *per capita* birth rate, a , and death rate, b , of the host. It is also assumed that in the absence of infection the disease-free population grows logistically²¹ to an environmental carrying capacity, K , where $K = (a - b)/\beta$ (the constant β measures the severity of density-dependent constraints on population growth). The system of equations for susceptible x , infected y and immune z

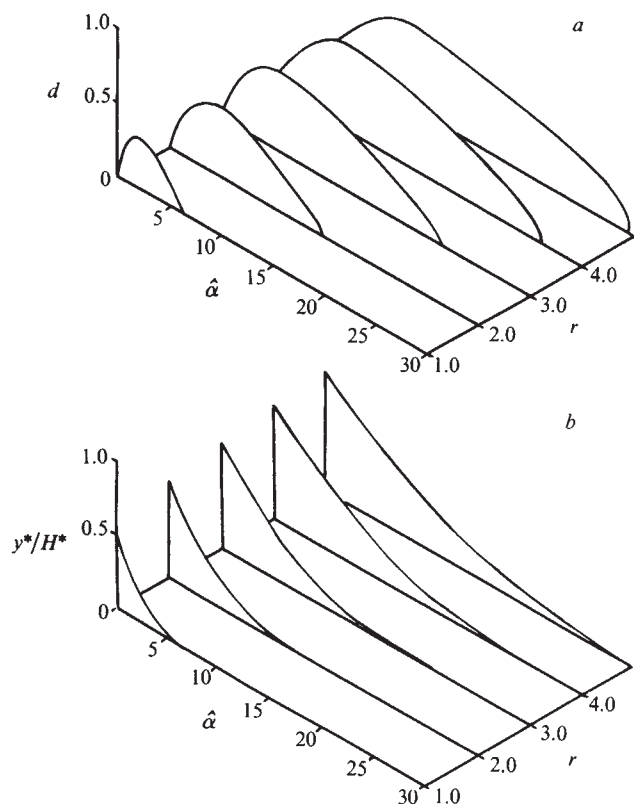


Fig. 1 Microparasitic infection model. Following conventional lines, the rate at which hosts acquire infection is assumed to be proportional to the number of encounters between susceptible and infected hosts, being Λxy , where Λ is a transmission coefficient²⁰. The mortality rate of infected hosts is taken to be $b + \hat{\alpha}$, with $\hat{\alpha}$ representing the mortality caused by the disease; there is also a recovery rate from infection, v . Recovered hosts are assumed to be initially immune, but this immunity can be lost at a rate γ (for permanent immunity $\gamma = 0$). These assumptions lead to the following equations: $dx/dt = aH - (b + \beta H)x - \Lambda xy + \gamma z$, $dy/dt = \Lambda xy - (b + \beta H)y - \hat{\alpha}y - vy$, $dz/dt = vy - (b + \beta H)z - \gamma z$. The construction of these equations follows conventional lines²⁰ (with the parameters a , b and β as defined in the main text) except for the inclusion of a *per capita* density-dependent host mortality rate $(b + \beta H)$ ²¹. The sum of the three equations for x , y and z yields: $dH/dt = (a - b - \beta H)H - \hat{\alpha}y$, equation (1). Provided $(a - b)$ is positive, the model yields locally stable equilibria. If host density, H , is less than a threshold value, H_T , where $H_T = (b + \hat{\alpha} + v)/(\Lambda - \beta)$, then after the introduction of infected hosts into the population, y will initially decrease and x increase. However, once x exceeds H_T , then y will increase and the system will converge (steadily or with damped oscillations) on one of two types of stable equilibrium state. Stable co-existence of host and parasite, where $H^* < K$ and $y^* > 0$, will result provided $\Lambda > \beta[1 + (b + \hat{\alpha} + v)/(a - b)]$. If this condition is not satisfied the parasite fails to co-exist and $H^* = K$ and $y^* = 0$. *a*, The influence of parasite pathogenicity, $\hat{\alpha}$, and the natural intrinsic growth rate of the host population r , where $r = a - b$, on the degree of host population depression d , where $d = 1 - H^*/K$. Parameter values of the model are as follows: $\beta = 0.01$, $\Lambda = 0.1$, $\gamma = 0.05$, $v = 0.5$. *b*, The influence of $\hat{\alpha}$ and r on the equilibrium proportion of infected hosts, y^*/H^* . Parameter values as for *a*.

hosts and their dynamical properties are detailed in the legend to Fig. 1. The model has two distinct patterns of population behaviour, namely (1) the disease persists and regulates the host population at a locally stable equilibrium H^* , where $H^* < K$, and $y^* > 0$; (2) the infection is not maintained within the host population and a stable disease-free equilibrium results where $H^* = K$ and $y^* = 0$.

The relationship between parasite pathogenicity, measured by $\hat{\alpha}$, the *per capita* rate of disease-induced mortality in the infected class of hosts, and the degree of depression of the host population equilibrium, d , where $d = 1 - H^*/K$ (varying

between 0 and 1), is shown in Fig. 1*a* for various values of r , the natural intrinsic growth rate of the host population ($r = a - b$). For a fixed value of r , the maximum degree of depression is achieved by parasites of moderate to low pathogenicity. Increased rates of parasite-induced host mortality result in a greater degree of host population depression until the rate of loss of infected hosts begins to have a detrimental effect on the efficiency of disease transmission. When the level of pathogenicity is very high, infected hosts die before effective transmission is achieved and the disease is thus unable to persist within the host population ($y^* \rightarrow 0$, $H^* \rightarrow K$). In brief, highly pathogenic organisms are likely to cause their own extinction but not that of their host. These conclusions withstand a variety of changes in model structure. Most importantly, if a totally immune category of hosts is absent, recovered hosts passing directly back to the susceptible class (as may often be the case with microparasitic infections of arthropods), a pattern emerges which is similar to that shown in Fig. 1.

Macroparasites with direct life cycles tend to produce persistent infections where the pathogenicity to the host, the rate of production of transmission stages and any resistance of the host to further infection all typically depend on the number of parasites present in a given host^{11,17-19,22}. A crude division of the host population into susceptible infected and immune classes is therefore inappropriate and a description of the dynamics must deal with the full probability distribution of parasites within the host population. Anderson and May⁹ have studied the dynamics of macroparasitic diseases with the aid of three differential equations for the number of hosts H , parasites P , and free-living infective stages W , which incorporate certain statistical moments of the distribution of parasite numbers per host. Parasite pathogenicity is measured by a parameter, α , which defines the disease induced *per capita* host death rate, per parasite, as a large amount of empirical evidence suggests that the rate of host mortality is proportional to parasite burden in a given host^{9,11,22}.

A model based on the work of Anderson and May⁹ but incorporating the assumption that in the absence of infection ($P = 0$) the host population grows in a logistic manner as defined in equation (1) (the details of the model are given in the legend to Fig. 2), predicts a similar relationship between parasite pathogenicity, α , and host population depression, d (Fig. 2), to that derived for microparasitic infections (Fig. 1). A further similarity in the predictions of the two models is shown in Figs 1*b* and 2*b*. Increased pathogenicity results in a decrease in both the proportion of infected hosts (y^*/H^* , microparasitic infections) and the mean parasite burden (P^*/H^* , macroparasitic infections). Both models suggest that observed levels of prevalence and intensity of infection within host populations in natural habitats will closely reflect parasite pathogenicity, a prediction supported by empirical evidence²³.

In contrast to these similarities, the impact of high rates of host population growth in the macroparasitic infection model is to lessen the degree of depression relative to the disease-free equilibrium value K , whereas the exact converse is predicted by the microparasite model. The explanation of this paradox lies in the ability of microparasites to colonise a host by direct multiplication within the host. The rate of establishment of new colonies in uninfected hosts is therefore enhanced by high host birth rates which continually replenish the stock of susceptible individuals. High rates of establishment increase the proportion of infected hosts in the population and thus increase the degree of population depression. Experimental studies of viral and bacterial infections in mouse populations clearly illustrate this principle²⁴.

In contrast, the population growth of macroparasites, which do not tend to exhibit direct multiplication, is dependent on the gradual accumulation of new infections within individual hosts and not on the input of susceptible individuals to the host population. High host birth rates tend to offset the mortalities caused by infections and hence decrease the degree of host population depression. Substantial depression of host popu-

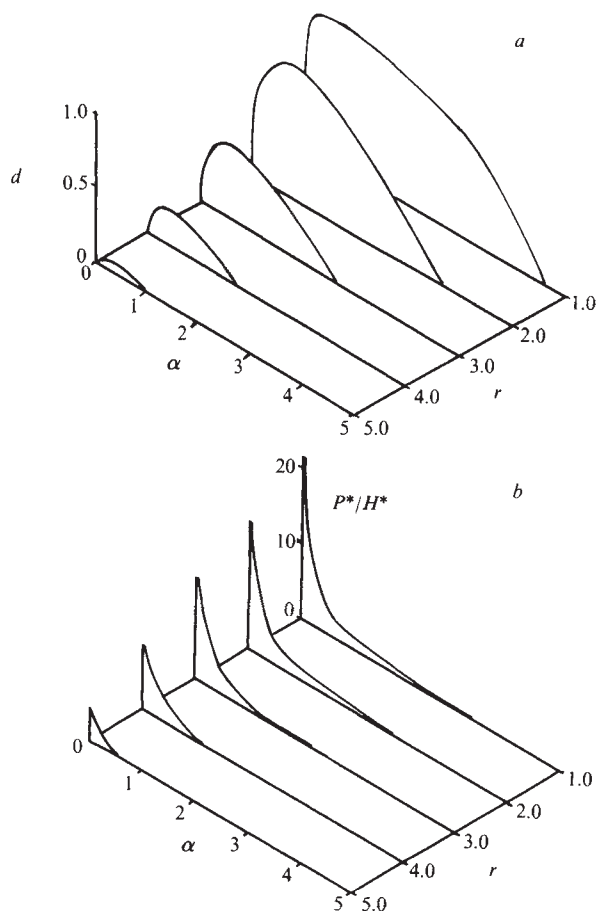


Fig. 2 Macroparasitic infection model. The framework of the macroparasite model is based on the work of Anderson and May^{9,11}, where the three equations for the population of hosts H , parasites P , and infective stages W are as follows: $dH/dt = (a - b - \beta H)H - \alpha P$, $dP/dt = \Lambda WH - (\mu + b + \alpha + \beta H)P - (\mu + \alpha)(k + 1)P^2/(kH)$, $dW/dt = \lambda P - cW - \Lambda WH$. The host birth and death rates a and b are as defined earlier for the microparasite model, as is the transmission parameter Λ (hosts acquire individual parasites at a rate proportional to the number of contacts between hosts and parasite infective stages, ΛWH) and the density-dependent host mortality rate, $b + \beta H$. The parasite-induced host death rate is assumed to be linearly proportional to the parasite burden in a given host and to occur at a rate α per parasite. The parasites are distributed as a negative binomial distribution with parameter k ; μ is the natural mortality rate of adult parasites, the net rate being dependent on the density of parasites within an individual host; λ is the rate of production of infective stages by an adult parasite, and c is the death rate of these infective stages. The empirical basis of these equations has been described in detail elsewhere⁹⁻¹¹. Provided $(a - b)$ is positive, the model yields locally stable equilibria. If host density H is less than a threshold value H_T , where $H_T = c(\mu + \alpha + \alpha)/[\Lambda(\lambda - \mu - \alpha - \alpha)]$, then the parasite cannot become established ($dP/dt < 0$). However, if $H > H_T$, provided $\lambda - (a + \mu + \alpha) > c\beta(a + \mu + \alpha)/[\Lambda(a - b)]$, the infection will become established and the equilibrium states $H^* < K$ and $P^* > 0$ are locally stable. If this condition is not satisfied the parasite fails to persist and $H^* = K$ and $P^* = 0$. *a*, The influence of parasite pathogenicity α and the natural intrinsic growth rate of the host population r on the degree of host population depression d . Parameter values of the model are as follows: $\beta = 0.01$, $\lambda = 8$, $\mu = 0.1$, $k = 1.0$, $c = 2.5$, $\Lambda = 0.5$. *b*, The influence of α and r on the equilibrium mean parasite burden per host, P^*/H^* . Parameter values as for *a*.

lations with high reproductive potentials may be achieved by macroparasites with the ability to produce large numbers of transmission stages to ensure the rapid accumulation of parasites within individual hosts.

The most important prediction of both models is related to the use of parasites as biological control agents; direct life-cycle

micro- and macroparasites of low to moderate pathogenicity are the most effective suppressors of host population growth. Very pathogenic organisms are likely to cause their own extinction but not that of their host. These conclusions seem to be reliable as they are derived independently from two mathematical models of very different structure.

Current practice is to select highly pathogenic organisms for the biological control of pest species¹⁻⁸. Such pathogens may cause high initial mortality within the pest population in a manner analogous to a single application of a chemical pesticide. However, in the absence of repeated introductions they are unlikely to achieve maximum and lasting depression of the pest population.

I thank Michael Hassel, John Lawton and Robert May for helpful discussions.

ROY M. ANDERSON

Department of Zoology and Applied Entomology,
Imperial College of Science and Technology,
London University,
Prince Consort Road,
London SW7, UK

Received 31 January; accepted 26 March 1979.

- Burges, H. D. & Hussey, N. W. (ed) *Microbial Control of Insects and Mites* (Academic, London, 1971).
- Huffaker, C. B. (ed.) *Biological Control* (Plenum, New York, 1974).
- DeBach, P. *Biological Control by Natural Enemies* (Cambridge University Press, 1974).
- McClelland, A. J. & Collins, P. *Nature* **276**, 548-549 (1978).
- Pramer, D. & Al-Rabaii, S. *Ann. N. Y. Acad. Sci.* **217**, 85-92 (1973).
- Chapman, H. C. A. *Rev. Ent.* **19**, 33-59 (1974).
- Aizawa, K. in *Microbial Control of Insects and Mites* (eds Burges, H. D. & Hussey, N. W.) 655-672 (Academic, London, 1971).
- Arata, A. A. *et al. Nature* **276**, 669-670 (1978).
- Anderson, R. M. & May, R. M. *J. Anim. Ecol.* **47**, 219-247 (1978).
- Anderson, R. M. in *Population Dynamics* (eds Anderson, R. M., Turner, B. D. & Taylor, L. R.) 245-281 (Blackwell, Oxford, 1979).
- May, R. M. & Anderson, R. M. *J. Anim. Ecol.* **47**, 249-268 (1978).
- Anderson, R. M. in *Some Mathematical Questions in Biology* Vol. 10 (ed. Levin, S. A.) (American Mathematical Society, Providence, 1979).
- Fenner, F., McAuslan, B. R., Mims, C. A., Sambrook, J. & White, D. O. *The Biology of Animal Viruses* (Academic, London, 1974).
- Noble, E. R. & Noble, G. A. *Parasitology: The Biology of Animal Parasites* (Lea and Febiger, Philadelphia, 1971).
- Mims, C. A. *The Pathogenesis of Infectious Diseases* (Academic, London, 1977).
- Meulen, V. & Hall, W. W. *J. gen. Virol.* **41**, 1-25 (1978).
- Wakelin, D. J. *Nature* **273**, 617-620 (1978).
- Soulsby, E. J. L. (ed.) *Immunity to Animal Parasites* (Academic, London, 1972).
- Van den Bossche, H. (ed.) *Biochemistry of Parasites and Host-Parasite Relationships* (North-Holland, Amsterdam, 1976).
- Bailey, N. T. J. *The Mathematical Theory of Infectious Diseases* (Griffin, London, 1975).
- Verhulst, P. F. *Corresp. Math. Phys.* **10**, 113-121 (1838).
- Anderson, R. M. *Parasitology* **76**, 119-157 (1978).
- Anderson, R. M. & May, R. M. *Parasitology* (in the press).
- Greenwood, M., Bradford Hill, A., Topley, N. W. C. & Wilson, J. *MRC Spec. Rep. No. 209* (HMSO, London, 1936).

The Stiles-Crawford hue shift following photopigment depletion

THE colour of a light is determined by such factors as its dominant wavelength, luminance, colorimetric purity, locus of retinal stimulation, surround chromaticity and preceding adapting field. Stiles and Crawford¹ were first to note that the position at which a small pencil of light enters the pupil also affects the light's colour. In general, light that enters through the centre of the pupil is seen as brighter than the same light impinging on the same retinal area but passing through the periphery of the pupil. This is known as the Stiles-Crawford effect of the first kind (S-C I). It has been demonstrated that this effect is due to the directional sensitivity of the receptors²⁻⁴ and is much more pronounced in cones than in rods⁵⁻⁸. Stiles⁹ also showed that two monochromatic lights of the same wavelength, one passing through the centre of the pupil (on-axis), the other entering through the periphery (off-axis), will differ in hue even after they are equated for brightness. (There is also a small change in

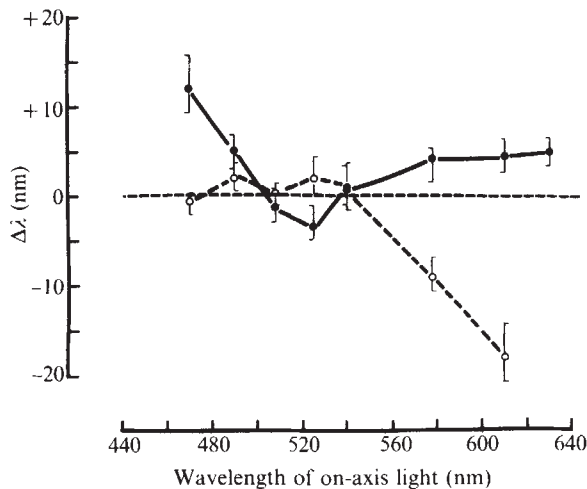


Fig. 1 S-C II for subject BW. $\Delta\lambda$ represents the shift required to bring an off-axis light to the same wavelength as an on-axis light after both lights have been matched in hue and brightness. Circles are means of two or three separately determined points of subjective equality. Vertical lines are the 5-95% range estimated from the psychometric functions used in deriving the means. ●, No-bleach condition; ○, high-bleach condition. The horizontal dashed line is the predicted result for the high-bleach condition based on the self-screening model.

saturation that becomes almost negligible at the long-wavelength end of the spectrum.) This is known as the Stiles-Crawford effect of the second kind (S-C II). For example, to match in hue an on-axis light of 610 nm an observer might require an off-axis light of 605 nm. The direction and magnitude of this hue shift varies with the wavelength of the on-axis light (Fig. 1). We have measured S-C II in conditions of varying levels of bleached photopigment. An explanation of the effect based on the absorbance of the photopigment can account for much but not all of the data. For the long-wavelength end of the spectrum, when the photopigment is significantly depleted, some other factor must have a role in mediating this effect.

Related to S-C II is the finding that the proportion of mixed red and green light required to match a yellow light (a metameric match) is different depending on where in the pupil the lights enter¹⁰. This upset in the metameric match must reflect the change in the shape of the spectral sensitivity curve of at least one of the three cone mechanisms when the light-entry position is changed. An altered shape in the spectral sensitivity curve of at least one of the cone systems can also account for S-C II.

Two hypotheses, that differ in their model of how light interacts with photoreceptors, have been proposed to explain these wavelength-dependent Stiles-Crawford effects. In so doing, these hypotheses have addressed the question of what it is that determines the spectral sensitivity of the cones, which is the fundamental property of the visual system on which our colour perception is based.

The waveguide hypothesis¹¹⁻¹³, based in part on physical optics, proposes that as the diameters of the photoreceptors are only slightly greater than the wavelength of light, the receptors act as dielectric (optical) waveguides. Light is propagated through and around an optical waveguide in non-uniform patterns of energy known as modal patterns. The amount of light travelling inside the guide is given in part by the kind of modal pattern produced. Also the amount of light available for absorption by the photopigment varies with the wavelength of the incident light and the angle of incidence. These, of course, are the two independent variables associated with S-C II. It is possible to derive a theoretical curve that closely matches the one seen connecting the filled circles in Fig. 1 by varying the parameters of the waveguide that determine its sensitivity (for example, refractive indices of the receptor and its surrounding medium, outer segment diameter, length and shape)^{14,15}.

According to the self-screening hypothesis¹⁶⁻¹⁸, which is based solely on geometric optics, light that enters through the pupil off-axis passes through less photopigment than on-axis light because of the former's oblique angle of incidence at the receptor layer. Such a shortening of the pathlength that the light takes through the pigment is equivalent to a reduction in the absorbance of the pigment. This follows from the Beer-Lambert law as does the fact that such a reduction has the effect of narrowing the spectral sensitivity curve of the cone system(s) involved. Self-screening has been shown to provide a reasonably good quantitative account of S-C II, if it is assumed that the absorbance of the photopigment(s) is about 0.5 or greater¹⁸. Based on direct^{19,20} and indirect¹⁷ estimates of the absorbance of human cone photopigments coupled with the assumption that passing light through the periphery of the pupil effectively shortens the pathlength, self-screening must to some extent be occurring and therefore must contribute at least in part to S-C II. Furthermore, self-screening provides a model with fewer assumptions and fewer free parameters than the waveguide model. We thus decided to test the generality of the self-screening model to see if S-C II could be accounted for entirely in terms of geometric optics and the absorption properties of photopigments.

In accord with the Beer-Lambert law, the lower the initial absorbance of a pigment, the smaller will be the range over which the absorbance can further be reduced. Thus, inherent in the self-screening model is the prediction that the magnitude of S-C II will be directly proportional to the initial absorbance of the photopigment. In the extreme case, when nearly all the photopigment is depleted, no measurable effect should be obtained.

It is possible to manipulate the initial absorbance of the photopigment by exposing the eye to an adapting light of a certain intensity. Such a bleaching light renders a given amount of pigment incapable of any further absorption of light. The reduction in concentration of available quantum-catching pigment thereby reduces the absorbance of the photopigment.

In our experiment, we measured S-C II in a no-bleach and a high-bleach condition in three subjects. Each subject was presented with two juxtaposed, circular fields each 20' in visual angle. The stimuli were presented in maxwellian view to one eye and were seen foveally. The source images, focused in the plane of the pupil, were 1 mm in diameter. One field, the standard, was

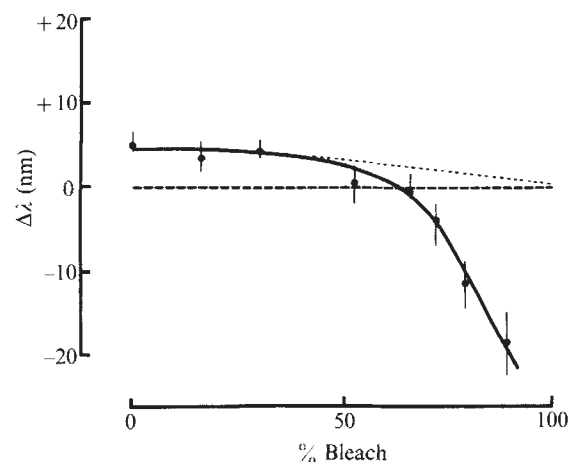


Fig. 2 S-C II at 610 nm as a function of the level of bleached photopigment for subject BW. The points are derived as described in Fig. 1, except that each point represents the result from a single session. Lightly dashed line shows the predicted result for bleaching levels above 40% based on the self-screening model. Below 40% the self-screening prediction and the obtained results coincide. Predictions are based on the receptor primaries of Vos and Walraven²⁴, assuming an absorbance of 0.5 for each pigment. It is also assumed that both pigments are bleached at the same rate by the white adapting field.

fixed in wavelength for a given session and entered through the pupil on-axis. The other field was varied around the standard in steps of 2 nm and entered through the pupil off-axis. (The position in the pupil that yields the peak luminous efficiency of a light is not always the centre. Therefore, we measured S-C I along the horizontal meridian to determine the pupillary position for each subject that represented the peak of his S-C I curve. For all subjects, the peak occurred within 1 mm of the geometric centre of the pupil. This was termed the on-axis condition. The off-axis condition was achieved by displacing the source image 3 mm to the opposite side of the pupil along the horizontal meridian.) The subject was required to state whether the variable field appeared to be a hue associated with a wavelength longer or shorter than the wavelength associated with the hue of the standard. Before each judgment, the subject matched the two fields in brightness. Residual differences in saturation, which were for the most part small, were disregarded. The luminances of the test fields were set by the subject to the lowest level possible in making a perfect match. Each wavelength chosen for the variable field was repeated five times. From the resulting psychometric function, the point of subjective equality could be determined. This was done for eight standards: 470, 490, 508, 525, 540, 578, 610 and 630 nm. All lights, except the white adapting field, were monochromatic (Bausch and Lomb grating monochromators, focal length, 0.5 m) with half-band widths of 3 nm. Each standard was repeated in a second session. If the results from the first and second session differed by more than 3 nm, a third session was included. For the high-bleach condition, following a 4-min exposure to a 5°, xenon-white adapting field that was calculated to bleach 88% of the mid- and long-wavelength sensitive cone pigments, each subject was presented with the following sequence of stimuli: a 2-s dark interval, a 1-s exposure of the test fields, a 10-s presentation of the bleaching field. This cycle, which maintained the percentage of pigment bleached at a constant level, continued for the duration of the session (about 0.5 h). The percentage of pigment bleached was calculated from Rushton and Henry's reflection densitometry measurements and pigment kinetic equations²¹. For the no-bleach condition, the subjects dark-adapted for 10 min and were then presented the test fields for 1 s every 4 s.

Figure 1 (filled circles) shows results for one subject for the no-bleach condition and are similar to the findings of others^{9,16,18}. The other subjects' data agree with those shown here. For example, when an on-axis light is set to 610 nm, the subject requires that an off-axis beam of light be approximately 605 nm for the two lights to match. Figure 1 (unfilled circles) are the results for the high-bleach condition. For this condition, the self-screening prediction (no hue shift) is represented by the horizontal, dashed line. While the data roughly agree with the prediction up to 540 nm, at longer wavelengths the empirical findings depart significantly from the theoretical curve. With 610 nm as an example again, in the no-bleach condition, an off-axis light of this wavelength appeared too red compared with an on-axis light of the same wavelength. In the high-bleach condition, a 610-nm, off-axis light appeared far too yellow when compared with a 610-nm, on-axis light.

We examined 610 nm at intermediate bleaching levels (Fig. 2). Self-screening predicts that as the percentage of bleached pigment increases, the magnitude of the effect should decrease until, finally, at close to a 100% bleach, the effect should be unmeasurable. The data are in agreement with the self-screening prediction for low and intermediate bleaching levels, but, above a 65% bleach, theory and data are discrepant. Our results for the long-wavelength end of the spectrum suggest that at high bleaching levels a factor other than self-screening, such as waveguide effects, must be invoked to account for S-C II. This is consistent with Weale's²² conclusion that long-wavelength light is treated preferentially in terms of the physical optics of the receptors. It might be that lights used to bleach a substantial portion of photopigment alter receptor properties in such a way as to increase a waveguide effect.

In conclusion, the present data indicate that a self-screening model can adequately account for S-C II when the eye is adapted to lights that bleach up to about 65% of the photopigment. This means that in such light-adapting conditions, the spectral sensitivity of each of the individual cone mechanisms is determined largely or solely by the absorption characteristics of their respective photopigments and the geometric optical properties of the receptors. However, when the photopigment is substantially depleted, possible waveguide effects of the receptors are revealed and have a significant role in determining spectral sensitivity.

This work was supported in part by an NIH National Research Service Award (5 F32 EY 05145) from the National Eye Institute to K.F. and a National Eye Institute grant (2 RO1 EY 01512) to B.R.W. Some of the data reported here appear in ref. 23.

KENNETH FULD
B. R. WOOTEN
LENORE KATZ

Walter S. Hunter Laboratory of Psychology,
Brown University,
Providence, Rhode Island 02912

Received 13 November 1978; accepted 19 March 1979.

1. Stiles, W. S. & Crawford, B. H. *Proc. R. Soc. B* **112**, 428-450 (1933).
2. Goldmann, H. *Ophthalmologica* **103**, 225 (1942).
3. Fankhauser, F., Enoch, J. M. & Cibis, P. M. *J. Ophthalmol.* **52**, 767-783 (1961).
4. Ripps, H. & Weale, R. A. *J. Physiol., Lond.* **173**, 57-64 (1964).
5. Crawford, B. H. *Proc. R. Soc. B* **124**, 81-96 (1937).
6. Stiles, W. S. *Proc. R. Soc. B* **127**, 64-105 (1939).
7. Flamant, F. & Stiles, W. S. *J. Physiol., Lond.* **107**, 187-202 (1948).
8. Van Loo, J. & Enoch, J. M. *Vision Res.* **15**, 1005-1009 (1975).
9. Stiles, W. S. *Proc. R. Soc. B* **123**, 90-118 (1937).
10. Brindley, G. S. *J. Physiol., Lond.* **122**, 332-350 (1953).
11. Toraldo di Francia, G. *J. opt. Soc. Am.* **39**, 324 (1949).
12. Enoch, J. M. *J. opt. Soc. Am.* **53**, 71-85 (1963).
13. Snyder, A. W. & Pask, C. *Vision Res.* **13**, 1115-1136 (1973).
14. Pask, C. & Snyder, A. W. in *Photoreceptor Optics* (eds Snyder, A. W. & Menzel, R.) 145-158 (Springer, Berlin, 1975).
15. Wijngaard, W. & van Kruysbergen, J. in *Photoreceptor Optics* (eds Snyder, A. W. & Menzel, R.) 175-183 (Springer, Berlin, 1975).
16. Enoch, J. M. & Stiles, W. S. *Optica Acta* **8**, 329-358 (1961).
17. Miller, S. S. *J. Physiol., Lond.* **223**, 89-107 (1972).
18. Walraven, P. L. & Bouman, M. A. *J. opt. Soc. Am.* **50**, 780-784 (1960).
19. Dobelle, W. H., Marks, W. B. & MacNichol, E. F. *Science* **166**, 1508-1510 (1969).
20. King-Smith, P. E. *J. Physiol., Lond.* **218**, 101-102P (1971).
21. Rushton, W. A. H. & Henry, G. H. *Vision Res.* **8**, 617-631 (1968).
22. Weale, R. A. *Vision Res.* **16**, 1395-1399 (1976).
23. Wooten, B. R., Fuld, K., Moore, M. & Katz, L. in *Visual Psychophysics and Physiology* (eds Armington, J. C., Krasukopf, J. & Wooten, B. R.) 245-256 (Academic, New York, 1978).
24. Vos, J. J. & Walraven, P. L. *Vision Res.* **11**, 799-818 (1971).

Estimating maximum limits to mutagenic potency from cytotoxic potency

RAPID and reliable screening methods are required for identifying environmental mutagens and estimating their mutagenic potency in preparation for use of more elaborate tests to assess the genetic risk to man. Many compounds are effectively screened by *in vitro* mutagenesis assays using either bacterial or mammalian cells. Although tests with mammalian cells are not as rapid and inexpensive as microbial assays, they are needed to confirm and extend the bacterial results, and probably provide a better assessment of potential risk to humans than do microbial tests. However, the screening of all potential mutagens with mammalian cell mutagenicity assays, although desirable, does not seem feasible unless, as we propose, test compounds are first pre-screened on the basis of their cytotoxic potency. On theoretical grounds, one expects a certain minimum cytotoxic potency to correlate with mutagenic potency, particularly when the latter is measured using forward mutations that result in inactive gene products. Specifically, cells cannot divide unless a substantial fraction of their genome is functionally intact; mutagenic agents should thus be obligatory cytotoxic agents, with a given mutagenicity conferring a certain irreducible cyto-

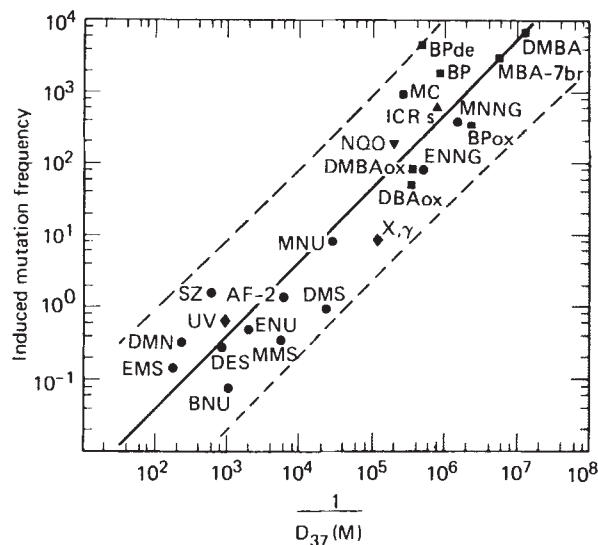


Fig. 1 The relationship between induced mutation frequency per viable cell (MF) per unit exposure dose (molar) and the reciprocal of D_{37} (M), that is, the dose required to kill 63% of the initial cell population. Killing is defined as the inability of a cell to undergo continued cell division *in vitro* resulting in a viable cell colony. The D_{37} values were estimated from (1) authors' precise survival responses; (2) computer fits in ref. 14; (3) published tabular data (occasionally over limited dose ranges). In some cases, the values are approximate but individual estimates generally vary no more than $\pm 20\%$ and never more than $\pm 50\%$ (in two cases). Mutation frequency data usually included at least two dose points in the most linear range, except for a few reports having only limited data where a single dose and zero dose (spontaneous frequency) were used. The data included in this plot are those for mutations at the *hprt* locus (hypoxanthine-guanine phosphoribosyltransferase, EC 2.4.2.7), the *aprt* locus (adenine phosphoribosyltransferase, EC 2.4.2.8), and the *tk* locus (thymidine kinase, EC 2.7.1.75). Results in Chinese hamster ovary cells^{7,14} ($n = 25$), V79 hamster cells¹⁵⁻²⁷ ($n = 18$), L5178Y mouse lymphoma cells²⁸⁻³¹ ($n = 4$), normal human diploid fibroblasts³²⁻³⁸ ($n = 9$) and human lymphoblasts^{39,40} ($n = 4$) were combined and expressed as the induced frequency of mutants resistant to azaguanine, AG^r ($n = 27$); thioguanine, TG^r ($n = 27$); azaadenine, AA^r ($n = 4$); bromodeoxyuridine, BUdR^r ($n = 2$). Abbreviations for the mutagens are given in Table 1. The linear regression line is given by the equation $(\log MF/M) = 1.03 (-\log D_{37}) - 3.49$, with a correlation coefficient of 0.93. The dotted lines represent boundaries of the 95% confidence band for Y data points estimated from X. The physical mutagens were entered on the figure after calculating their potencies relative to EMS, based on the ratios of the number of lesions induced in the DNA of a cell at the D_{37} doses (UV-induced dimers⁴¹ and total ionisation lesions^{5,6}) to the number of alkylations⁴² induced by EMS at its D_{37} . The 'molar equivalent D_{37} ' values were obtained by multiplying the D_{37} for EMS (6×10^{-3} M) by the potency ratios for UV (0.18) and ionising radiation (1.4×10^{-3}). The 'molar equivalent mutation frequencies' were obtained by multiplying the geometric means of the observed mutation frequencies per D_{37} for UV and ionising radiation by the respective molar equivalent D_{37} values. The calculated parameters were added to the graph, but were not included in the regression calculation. The lower detection limit of the assays (stippled area) was estimated, assuming that (1) at frequencies below 10^{-5} induced mutations per viable cell, statistically significant determinations of MF are generally not practical, and (2) the maximum feasible exposure dose is usually ≤ 10 -fold higher than the D_{37} concentration.

toxicity. We show here that the cytotoxic potency of 22 chemical mutagens is highly correlated with their mutagenic potency as assayed in five rodent and human *in vitro* cell systems. This relationship implies that the maximum potential mutagenic potency of such compounds may be reliably estimated from rapid and straightforward measurements of their cytotoxic potency, the latter defined as the failure of cultured cells to undergo continued cell division. The estimate is necessarily a maximum one, as an agent may exert cytotoxic effects by pathways independent of its mutagenic action.

A correlation between cell survival and mutagenic response has been noted for ionising radiation mutagenesis¹ and also seems to be applicable to ultraviolet radiation and certain chemicals²⁻⁴. Roberts *et al.* have discussed the molecular aspects of cellular lethality and mutagenesis in terms of damage to DNA^{5,6}. To test this relationship further, we have compiled data on the cytotoxicity and mutagenicity of the 24 chemical and physical mutagens listed in Table 1. As the measure of cytotoxic

potency, we have used the D_{37} unit of survival, defined here as the dose required to kill approximately 63% of the initial cell population. Unlike D_{01} , estimation of D_{37} is usually possible from the limited survival data that accompany published mutation frequencies; D_{37} generally measures survival within the linear portion of the dose response for mutation (lower survivals are frequently out of the linear mutation range). Moreover, this parameter reflects differences in repair capability or fidelity that may influence both toxicity and mutagenicity.

Figure 1 compares the frequency of forward mutations induced per applied dose with the reciprocal of the molar dose required for D_{37} survival in three rodent and two human cell systems⁵⁻⁴². As plotted, the data constitute a potency index, extending from weakly mutagenic and less toxic compounds to highly mutagenic agents that are toxic at very low concentrations. The relative increase in cytotoxicity is accompanied by a proportional increase in mutagenicity, that is, the slope of the regression relationship is 1.03 ± 0.03 . The observed correlation between cytotoxicity and mutagenicity suggests that a determination of the D_{37} dose of toxic agents can be used to estimate the maximum potential-induced mutation frequency. In particular, these data permit the hypothesis that no agents will be found whose location on this graph is significantly above the upper dotted line. Whereas mutagenicity and cytotoxicity separately range over six orders of magnitude, no agent can

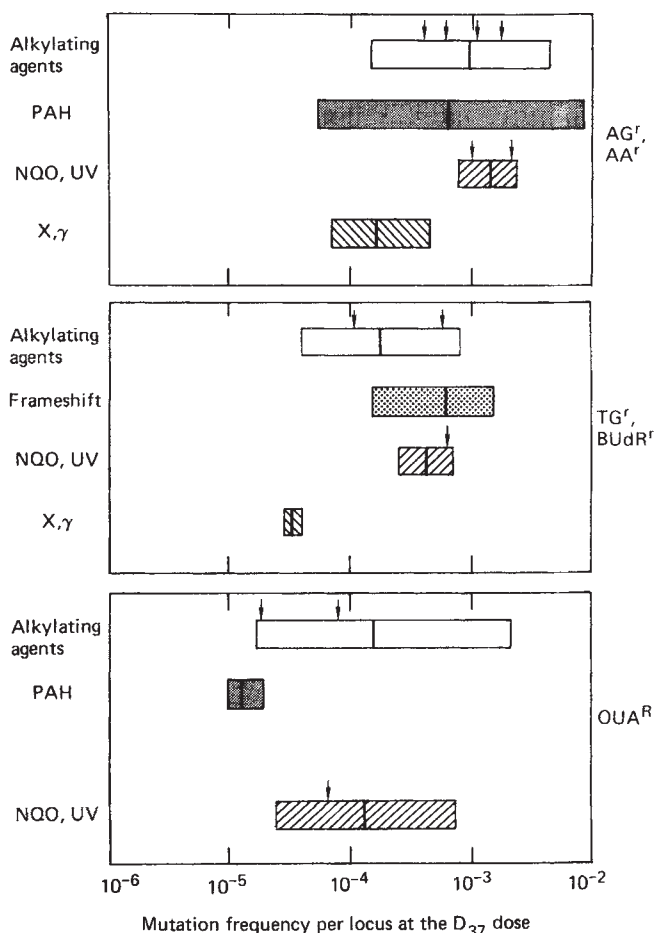


Fig. 2 The induced frequency per viable cell (MF) calculated for the D_{37} exposure dose of the physical and chemical mutagens listed in Table 1. The loci include *hprt* (AG^r, TG^r), *aprt* (AA^r) and *tk* (BUdR^r), as well as resistance to ouabain (OUA^r) involving presumed lesions in the Na⁺-K⁺-ATPase enzyme (EC 3.6.1.3). The inner mark within the bars represents the geometric mean of data for all systems, with the range of reported values indicated by the ends of each bar. The arrows refer to the author's data for the Chinese hamster ovary multiple marker system⁷. According to the assumptions listed in Fig. 1, the lower limit for determining the MF at the D_{37} would be 10^{-6} .

Table 1 Compounds evaluated for mutagenic potency

Compound	n	Ref.
Alkylating agents		
Streptozotocin (SZ)	1	17
Dimethylnitrosamine (DMN)	2	23, 31
Ethyl methanesulphonate (EMS)	15	7, 11–13, 15, 16, 26, 28, 29, 35, 42
Diethylsulphonate (DES)	1	10
N-methyl-N'-nitro-N-nitrosoguanidine (MNNG)	13	7, 9, 13, 17, 18, 21, 25, 27, 34, 39, 40
N-methyl-N-nitrosourea (MNU)	2	9, 40
N-ethyl-N-nitrosourea (ENU)	1	9
2-(2-furyl)-3-(5-nitro-2-furyl)acrylamide (AF-2)	2	27, 36
N-ethyl-N'-nitro-N-nitrosoguanidine (ENNG)	1	9
N-butyl-N-nitrosourea (BNU)	1	9
Methyl methanesulphonate (MMS)	2	10, 30
Dimethylsulphate (DMS)	1	10
Polycyclic aromatic hydrocarbons (PAH)		
Benzo(a)pyrene, diol epoxide (BPde)	1	24
3-Methylcholanthrene (MC)	2	22
Benzo(a)pyrene (BP)	3	20, 22
7,12-Dimethylbenz(a)anthracene (DMBA)	2	22
7-Bromomethylbenz(a)anthracene (MBA-7br)	1	19
7,12-Dimethylbenz(a)anthracene, 5,6-oxide (DMBAox)	1	37
Dibenz(a,h)anthracene, 5,6-oxide (DBAox)	1	37
Benzo(a)pyrene, 4,5-oxide (BPox)	1	37
Frameshift mutagens		
Heterocyclic nitrogen mustard compounds (ICRs)	5	14, 40
Physical mutagens		
X and γ rays	4	2, 6, 8, 32, 33
Ultraviolet irradiation (UV)	6	11, 13, 16, 26, 38, 41
Miscellaneous		
4-Nitroquinoline-1-oxide (NQO)	4	7

n, Number of determinations evaluated for each compound (some at multiple genetic loci).

induce more than approximately one forward mutation at a given locus per 100 cells surviving at the D₃₇ dose. If this hypothesis is valid, this is a biological limit to mutagenic potency that is demonstrated by compounds whose cytotoxicity is due solely to mutational events.

On the other hand, we believe the absence of points below the lower dotted line to be artefactual, representing the bias of reported mutagenesis studies for compounds whose ratio of mutagenicity to cytotoxicity is relatively high. There are compounds which would graph somewhere in this lower range; for example, a few derivatives of heterocyclic nitrogen mustard frameshift mutagens¹⁴, a metabolite of a polycyclic hydrocarbon³⁷, and an inhibitor of DNA synthesis⁴³ are highly toxic, but are apparently non-mutagenic in mammalian *in vitro* test systems. Also, certain chemicals yield multi-phasic cell survival curves that cast doubt on the validity of the D₃₇ concept for these mutagens. The shaded area in Fig. 1 reflects values below the detection limit for the systems studied. The cytotoxicity data for non-mutagens would also fall within this range. Thus, the dynamic range for forward mutations in mammalian *in vitro* systems is approximately four orders of magnitude (from the lowest detectable level to the highest possible mutagenic potency at the D₃₇ level of cytotoxicity). Compounds near the bottom of this range will presumably be of greater concern as cytotoxic agents than as mutagens.

The relative mutagenicity of different compounds at a given cytotoxicity (Fig. 1) is not significantly related to the class of mutagen or the forward mutation marker scored. This is shown in Fig. 2, which displays the observed values of the mutation frequency per locus at the D₃₇ dose for different classes of mutagens at different genetic loci. This variation is also not due to systematic differences among cell systems (data not shown).

Cytotoxicity, as defined (Fig. 1), may result from genetic (DNA-related) and non-genetic mechanisms of injury. Agents that induce the highest mutation frequencies at the D₃₇ dose may have the tightest coupling between genetic injury and lethality, that is, the cytotoxicity related to DNA damage pre-

dominates over other forms of general toxicity. Agents inducing lower mutation frequencies at the D₃₇ dose may produce a larger proportion of non-genetic injury.

The remarkable similarity among mutation frequencies at equitoxic concentrations has been noted by others for several mutagens^{5,6,17,37}. The correlation between cytotoxicity and mutagenicity need not be due solely to the same type of cellular damage causing both lethal and mutational events⁴; factors that influence the effective exposure of the cells to chemical mutagens (toxicification, detoxification) may affect both end points proportionally.

Additional data are needed to establish that partitioning of genetic and non-genetic toxicity induced by a mutagenic agent can be estimated from the magnitude of mutation induced at the D₃₇ dose. However, the data in Fig. 1 indicate that cytotoxicity as measured by the D₃₇ dosage provides a quantitative estimate of potential mutagenicity. Where appropriate, more specific bioassays can then be applied to determine precisely the *in vitro* mutagenic potency.

This work was performed under the auspices of the US Department of Energy by the Lawrence Livermore Laboratory under contract no. W-7405-ENG-48 and Interagency Agreements IAG-D5-E681-AN and AO with the Environmental Protection Agency.

J. H. CARVER
F. T. HATCH
E. W. BRANSCOMB

Lawrence Livermore Laboratory,
University of California,
PO Box 5507,
Biomedical Sciences Division,
Livermore, California 94550

Received 29 November 1978; accepted 5 March 1979.

- Thacker, J., Stretch, A. & Stephens, M. A. *Mutat. Res.* **42**, 313–326 (1977).
- Chadwick, K. H., Leenhouts, H. P., Szumiel, I. & Nias, A. H. W. *Int. J. Radiat. Biol.* **30**, 511–524 (1976).
- Parodi, S. & Brambilla, G. *Mutat. Res.* **47**, 53–74 (1977).
- Munson, R. J. & Goodhead, D. T. *Mutat. Res.* **42**, 145–160 (1977).
- Roberts, J. J. in *Advances in Radiation Biology* (eds Lett, J. T. & Adler, H.) 212–436 (Academic, New York, 1978).
- Roberts, J. J., Sturrock, J. E. & Ward, K. N. in *Chemical Carcinogenesis Part A* (eds Ts'o, P. O. & Dipple, J. A.) 401–425 (Marcel Dekker, New York, 1974).
- Carver, J. H., Wanders, D. L., Adair, G. M. & Branscomb, E. W. *Mutat. Res.* **53**, 96–97 (1978).
- Carver, J. H., Dewey, W. C. & Hopwood, L. E. *Mutat. Res.* **34**, 465–480 (1976).
- Couch, D. B. & Hsie, A. W. *Mutat. Res.* **41**, 361–376 (1976).
- Couch, D. B., Forbes, N. L. & Hsie, A. W. *Mutat. Res.* **57**, 217–224 (1978).
- Hsie, A. W., Brimer, P. A., Mitchell, T. J. & Gosslee, D. S. *Somatic Cell Genet.* **1**, 247–261 (1975).
- Jones, G. E. & Sargent, P. A. *Cell* **2**, 43–54 (1974).
- Molnar, S. J. & Rauth, A. M. *Mutat. Res.* **41**, 361–376 (1976).
- O'Neill, J. P., Fuscoe, J. C. & Hsie, A. W. *Cancer Res.* **38**, 506–509 (1978).
- Abbondandolo, A. *et al. Mutat. Res.* **37**, 293–306 (1976).
- Arlett, C. F., Turnbull, D., Harcourt, S. A., Lehmann, A. R. & Colella, C. M. *Mutat. Res.* **33**, 261–278 (1975).
- Bhuyan, B. K., Peterson, A. R. & Heidelberger, C. *Chem.-Biol. Interactions* **13**, 173–179 (1976).
- Davies, P. J. & Parry, J. *Genet. Res.* **24**, 311–314 (1974).
- Duncan, M. E. & Brookes, P. *Mutat. Res.* **21**, 107–118 (1973).
- Huberman, E. *Mutat. Res.* **29**, 285–291 (1975).
- Huberman, E., Aspiras, L., Heidelberger, C., Grover, P. L. & Sims, P. *Proc. natn. Acad. Sci. U.S.A.* **68**, 3195–3199 (1971).
- Huberman, E. & Sachs, L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 188–192 (1976).
- Kuroki, T., Drevon, C. & Montesano, R. *Cancer Res.* **37**, 1044–1050 (1977).
- Newbold, R. F. & Brookes, P. *Nature* **261**, 52–54 (1976).
- Peterson, A. R. *et al. Mutat. Res.* **36**, 345–356 (1976).
- van Zeeland, A. A. & Simons, J. W. I. M. *Mutat. Res.* **35**, 129–138 (1976).
- Wild, D. *Mutat. Res.* **31**, 197–199 (1975).
- Clive, D., Flamm, W. G., Machesko, M. R. & Bernheim, N. J. *Mutat. Res.* **16**, 77–87 (1972).
- Cole, J. & Arlett, C. F. *Mutat. Res.* **34**, 507–526 (1976).
- Cole, J. & Arlett, C. F. *Mutat. Res.* **50**, 111–120 (1978).
- Frantz, C. N. *J. Tox. envir. Hlth* **2**, 179–187 (1976).
- Albertini, R. J. & DeMars, R. *Mutat. Res.* **18**, 199–224 (1973).
- Cox, R. & Masson, W. K. *Mutat. Res.* **37**, 125–136 (1976).
- Jacobs, L. & DeMars, R. *Mutat. Res.* **53**, 29–53 (1978).
- Kuroda, Y. *Jap. J. Genet.* **49**, 389–398 (1974).
- Kuroda, Y. *Mutat. Res.* **30**, 229–238 (1975).
- Maher, V. M., McCormick, J. J., Grover, P. L. & Sims, P. *Mutat. Res.* **43**, 117–138 (1977).
- Maher, V. M., Ouellette, L. M., Curren, R. D. & McCormick, J. J. *Nature* **261**, 593–594 (1976).
- Penman, B. W. & Thilly, W. G. *Somatic Cell Genet.* **2**, 325–330 (1976).
- Thilly, W. G., Deluca, J. G., Hoppe, H. V. & Penman, B. W. *Mutat. Res.* **50**, 137–144 (1978).
- Williams, J. I. & Cleaver, J. E. *Biophys. J.* **22**, 265–279 (1978).
- Aaron, C. S., van Zeeland, A. A., Mohn, G. R. & Natarajan, A. T. *Mutat. Res.* **50**, 419–426 (1978).
- Bradley, M. O. & Sharkey, N. A. *Nature* **274**, 607–608 (1978).

Hyperdiploid species hybrids for gene mapping in *Xenopus*

SPONTANEOUS endo-reduplication in the female germ line of hybrids has made it possible to produce several strains of isogenic *Xenopus*, providing one of the many features required of a laboratory animal^{1,2}. The usefulness of *Xenopus* in the laboratory, however, would be improved if a gene map were available. We report here that hybrids can also be used to define linkage groups and their assignment to chromosomes. Our method is based on the possibility of creating aneuploid *Xenopus* which are diploid for the chromosomes of one species but have supernumerary chromosomes of a second species³. This method complements those already available⁴⁻⁸, and has the advantage that because traits can be monitored in the whole animal, a large repertory of markers will be made available.

Figure 1 is a schematic presentation of how such animals are obtained, in this case for *X. laevis* × *X. gilli* hybrids. First generation hybrid females from this and other species combinations frequently produce diploid eggs from allotetraploid oocytes produced by endo-reduplication in the germ line. These oocytes have a complete set of bivalents from each parent species, synapsis being restricted to endo-reduplicated identical chromosomes^{1,9}. Fertilised by sperm of one of the parent species, such eggs develop into triploid females (WZZ) because the female determinant W is dominant¹⁰. Triploid oocytes of such backcross females show a diploid number of bivalents, presumably the two homologous sets of chromosomes, with the third genome consisting of univalents. No trivalents have been observed⁹. The eggs thus contain a complete haploid chromosome set of the backcross species, *X. laevis* in our example, and, on the assumption of a random distribution of univalents, 0–18 chromosomes of the other species. Random distribution is indeed observed in the case of the easily recognisable nucleolus organiser chromosome. Fertilised again by *X. laevis* sperm, hyperdiploid embryos are obtained which are diploid for *X. laevis* but trisomic for a random selection of *X. gilli* chromosomes. Most of these zygotes undergo abnormal embryogenesis and die³, but 1–3% of them reach adulthood.

We have analysed 14 such hyperdiploid adults for karyotype and for the expression of *X. gilli*-specific characters. Antisera raised in *X. laevis* (as described elsewhere¹¹) enabled us to detect by immunoelectrophoresis up to six *X. gilli*-specific serum antigens (Fig. 2), one of which was the heavy chain of low-molecular weight immunoglobulin (7S, IgG-like). Antiserum B10 was made specific for *X. gilli* Ig by immunoabsorption; A84 also contained antibodies against *X. gilli* Ig. We have not established the homologies between the serum proteins revealed by the different antisera. Two *X. gilli*-specific erythrocyte antigens were distinguished by lysis after isoelectrofocusing of specific antibodies (antisera A38 and A39).

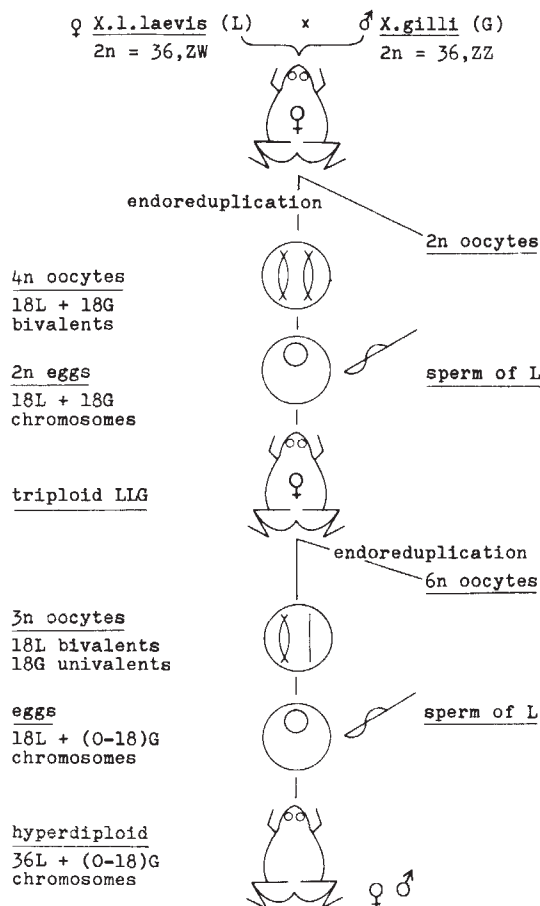


Fig. 1 Schematic presentation of pedigree of hyperdiploid *Xenopus* species hybrids.

Haemagglutination with antiserum A37 revealed an additional *X. gilli* erythrocyte antigen. Table 1 shows that *X. gilli* Ig and the erythrocyte surface antigens segregated as unlinked characters. They were not linked with either the single serum protein revealed by antiserum A85 or with those revealed by antiserum A82. This is in agreement with results obtained on diploid backcross animals of clone LG 15 (ref. 11) where, in contrast to most other LG clones, there was recombination between *X. gilli* and *X. laevis* chromosomes². These characters depend therefore on genes on separate chromosomes. The lack of linkage between erythrocyte surface antigens suggests that some of them may represent minor histocompatibility antigens rather than major ones.

Chromosomes were prepared from phytohaemagglutinin-stimulated leukocyte cultures¹². The karyotypes of both species are similar ($2n = 36$) except for a chromosome pair with a

Table 1 Number of supernumerary *X. gilli* chromosomes and occurrence of *X. gilli*-specific serum and erythrocyte antigens in a sample of hyperdiploid *Xenopus* hybrids

<i>X. gilli</i> specificities	Control LG	Hyperdiploid LL(G) animal no.													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Number of supernumerary chromosomes	(18)	11	13	16	8	8	8	14	8	8	.	11	8	8	8
Immunoglobulin heavy chain:															
Antiserum B10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Number of erythrocyte antigens:															
Agglutination, antiserum A37	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+
Lysis, antisera A38 + A39	2	2	2	2	2	2	2	2	1	2	2	0	1	1	2
Number of serum proteins:															
Antiserum A82	3	1	3	3	3	3	.	3	3	3	3	0	0	3	3
A83	6	6	6	5	5	5	.	6	6	6	4	0	1	2	4
A84	5	3	5	5	4	4	.	5	3	5	4	.	2	5	.
A85	1	0	1	1	1	1	.	1	1	0	0	.	0	1	.

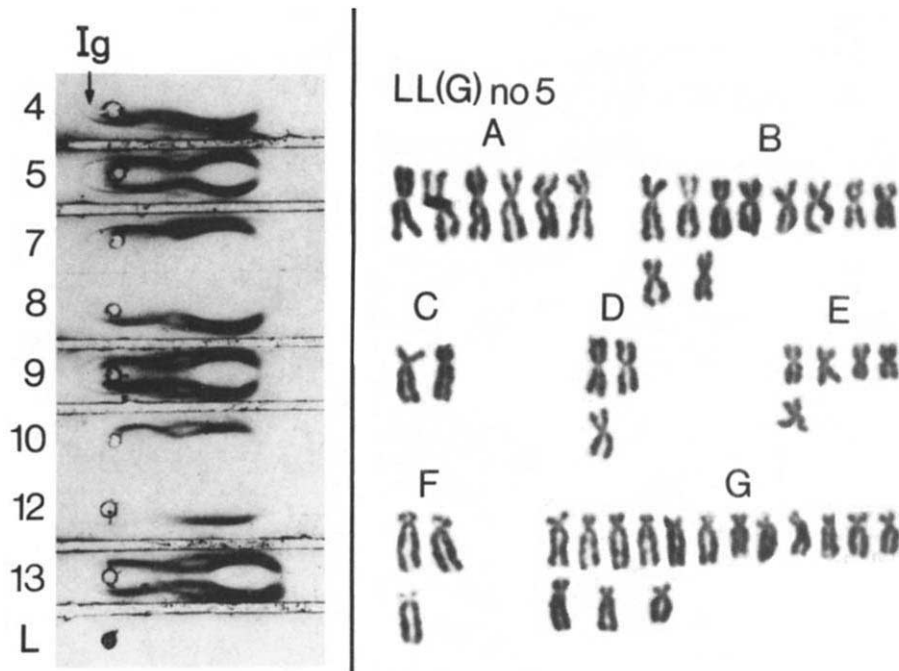


Fig. 2 Immunoelectrophoretic analysis of the serum of eight hyperdiploid *Xenopus* hybrids with the antiserum A84. Serum samples in the wells from LL(G) animals nos. 4–13, and *X. laevis* (L) control. Antiserum A84 in the trough. Cathode to the left. Migration conditions: 1 h, pH 8.6, 5 mA per slide. Note the expression of only two *X. gilli* antigens, one of which is immunoglobulin (Ig, heavy chain) by animal no. 12, and 3 to 5 antigens by the other specimens. *X. laevis* control is negative. The karyotype of animal no. 5 shows groups (A–G) of similar chromosomes; supernumerary chromosomes on a second line.

secondary constriction specific for *X. gilli*¹³. The chromosomes can be classified into seven groups (Fig. 2) on the basis of relative length and arm ratio⁶, but individual chromosome recognition requires differential staining. Whereas the total number of supernumerary chromosomes can be determined (Table 1), their assignment to groups remains uncertain. Nevertheless, the nucleolus organiser chromosome (group F) is easily recognised in both mitotic and interphase cells and can be excluded as a carrier of one of the characters checked so far. Acrocentric chromosomes can also be counted. Animals 9, 13 and 14 had only two supernumerary acrocentric chromosomes; thus, not more than two of the characters analysed may depend on group G chromosomes. The four specimens which were not trisomic for the chromosome of group D were also deficient for one of the *X. gilli* erythrocyte antigens, indicating the possible chromosome location of the corresponding gene.

Some animals had a hypertriploid constitution due to failure of second meiotic division. Because the second meiotic division segregates chromatids, eggs which retain the second polar body contain a diploid set of *X. laevis* chromosomes as well as both chromatids of the *X. gilli* univalents distributed at meiosis I. Zygotes are, therefore, triploid for *X. laevis* and supernumerary for pairs of *X. gilli* chromosomes. Animal 12 had a bimodal chromosome number with 36 and 44 chromosomes. It is not certain whether the karyotype found in a given tissue is representative for the whole animal. Mosaicism could arise after non-disjunction during ontogenesis because the original zygotic constitution might be improved through the loss of supernumerary chromosomes.

It is striking that in spite of the variable degree of hyperdiploidy, most animals had most of the *X. gilli* characters studied. Was this attributable to redundancy of genes? Because several levels of ploidy occur in *Xenopus*¹⁴, one might consider both *X. laevis* and *X. gilli* to be ancient tetraploids. Even a restricted number of supernumerary chromosomes could then contain one of the originally duplicated genes. This would also imply that activity of duplicated genes has been conserved during evolution and that our methods do not distinguish between the various alleles. It seems more likely that survival of aneuploids is better when animals are functionally hypotriploid rather than trisomic for only a few chromosomes.

HANS RUDOLF KOBEL

Génétique Animale et Végétale,
Université de Genève,
CH-1224 Geneva, Switzerland

LOUIS DU PASQUIER

Basel Institute for Immunology,
CH-4058 Basel, Switzerland

Received 19 January; accepted 15 March 1979.

1. Kobel, H. R. & Du Pasquier, L. *Immunogenetics* **2**, 87 (1975).
2. Kobel, H. R. & Du Pasquier, L. in *Developmental Immunobiology* (eds Solomon & Horton) 299 (North-Holland, Amsterdam, 1977).
3. Kobel, H. R. *Arch. Genetik* **51** (1978).
4. Pardue, M. L. *Cold Spring Harb. Symp. quant. Biol.* **38**, 475 (1973).
5. Pardue, M. L., Brown, D. D. & Birnstiel, M. L. *Chromosoma* **42**, 191 (1973).
6. Tymowska, J. & Kobel, H. R. *Cytogenet. Cell Genet.* **11**, 270 (1972).
7. Funaki, K., Matsui, S. & Sasaki, M. *Chromosoma* **49**, 357 (1975).
8. Schmid, M., Hofgärtner, F. J., Zenzes, M. T. & Engel, W. *Hum. Genet.* **38**, 279 (1977).
9. Müller, W. P. *Chromosoma* **59**, 273 (1977).
10. Chang, C. Y. & Witschi, E. *Proc. Soc. exp. Biol. Med.* **93**, 140 (1956).
11. Du Pasquier, L. & Kobel, H. R. *Immunogenetics* (in the press).
12. Du Pasquier, L. & Horton, J. D. *Immunogenetics* **3**, 105 (1976).
13. Tymowska, J. thesis, University of Geneva (1977).
14. Thiébaud, Ch. H. & Fischberg, M. *Chromosoma* **59**, 253 (1977).

17 β -Carboxamide steroids are a new class of glucocorticoid antagonists

GLUCOCORTICOID hormones exert many of their effects by modifying the activity of certain enzymes in target tissues, probably as a result of alterations in gene expression¹. Experimental models have been developed in which enzyme induction by various steroids correlates well with their anti-inflammatory, thymolytic and glycopexic properties. This is the case for tyrosine transaminase in cultured rat hepatoma (HTC) cells^{2,3}. Studies in this system showed that, in addition to glucocorticoid agonists, partial agonists and antagonists could be identified⁴. The latter compete with agonists for binding to the intracellular glucocorticoid receptor but do not trigger the glucocorticoid effect⁵. Compounds which specifically antagonise glucocorticoid action at the target cell would be useful for the treatment of certain diseases characterised by excessive production of glucocorticoids¹ and for further studying the molecular mechanism of action of these hormones. However, unlike anti-androgens and anti-oestrogens⁶, the only known glucocorticoid antagonists are natural steroid hormones (such as progesterone and testosterone) or their analogues, which interact with their own receptors and plasma binding proteins and are therefore not very

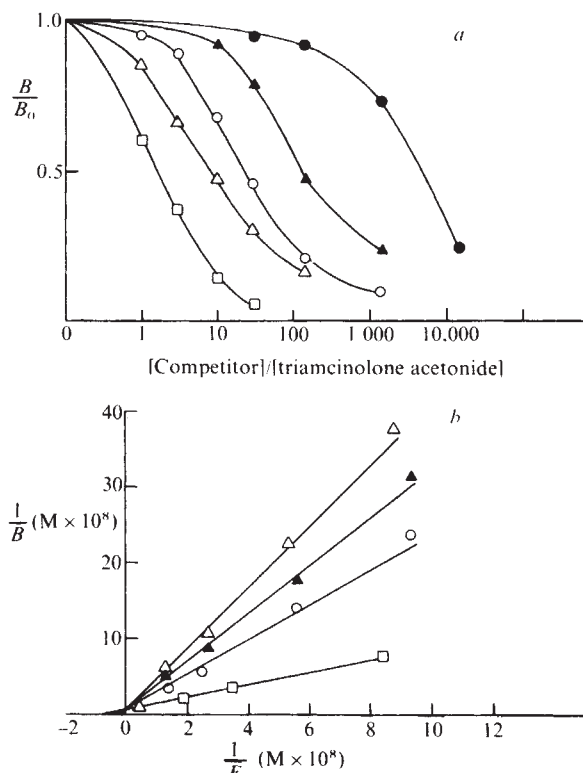


Fig. 1 Glucocorticoid receptor binding of 17 β -carboxamide derivatives of dexamethasone. Livers from adrenalectomised male rats (Wistar, 250 g) were homogenised in buffer containing 20 mM Tris, 20 mM 2-mercaptoethanol, 3 mM MgCl₂ and 20% glycerol, pH 7.5. The homogenate was centrifuged at 105,000g for 2 h to yield the cytosol. Cytosol aliquots were incubated with [1,2,4(n)-³H]triamcinolone acetonide (28 Ci mmol⁻¹, Radiochemical Centre). After 2 h at 0°C, specific binding to the glucocorticoid receptor was determined in duplicate by charcoal adsorption assay⁵. *a*, Inhibition of binding of 35 nM triamcinolone acetonide by increasing concentrations of dexamethasone and four of its derivatives (for nomenclature, see Table 1). Ordinate: ratio of binding of ³H-labelled triamcinolone acetonide in the presence of competitor (*B*) to binding in the absence of competitor (*B*₀). Abscissa: concentration ratios. □, Dexamethasone; △, DXB; ○, DXP; ▲, DXN; ●, DXO. *b*, Double-reciprocal plot of the specific binding (*B*) of triamcinolone acetonide at various concentrations (*F*), in the absence of competitor and the presence of a constant concentration of competitors. △, 0.25 μ M DXB; ▲, 1 μ M DXN; ○, 0.5 μ M DXP; □, no competitor.

specific. We report here that new derivatives of dexamethasone, a specific glucocorticoid agonist which does not bind to rat or human plasma transcortin⁵, can block the induction of tyrosine transaminase in HTC cells.

The novel steroids listed in Table 1 were synthesised and purified in our laboratory (P.F. and P.L., in preparation). They include the 17 β -carboxylic steroids obtained by periodic oxidation of cortisol, corticosterone and dexamethasone, respectively, and seven 17 β -carboxamide derivatives. The possible interaction of these steroids with the glucocorticoid receptor was studied in liver cytosol from adrenalectomised rats, using as specific ligands ³H-labelled dexamethasone or triamcinolone acetonide, both synthetic agonists. Even at a 1,000-fold excess the 17 β -carboxylic compounds DCAC, HCAC and DXO (see Table 1 for structures) barely inhibited binding of the ³H-labelled steroids. However, after coupling the 17 β -carboxyl group of DXO with various amines, clear inhibition by these dexamethasone derivatives was observed (Fig. 1*a*). The benzylamine-substituted compound DXB was the most active, a 10-fold excess being sufficient to produce a 50% reduction in the binding of the potent glucocorticoid, triamcinolone acetonide (Fig. 1*a*). This phenomenon was further investigated using different concentrations of radioactive ligand. Double-reciprocal plots suggested that the 17 β -carboxamide derivatives

inhibited binding of triamcinolone acetonide to a homogeneous population of receptor sites in a competitive manner. Representative experiments are shown in Fig. 1*b*. The order of affinity was DXB > DXH > DMP > DXP > DXM > DXN > DCACP > DXO, with an equilibrium dissociation constant for DXB (0.8–1.1 $\times 10^{-7}$ M) as high as that of corticosterone, the natural glucocorticoid in rats.

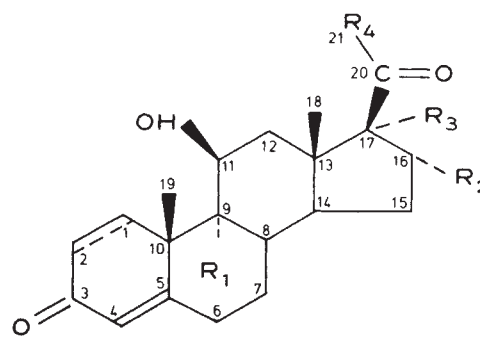
One might therefore predict that some of the dexamethasone derivatives would mimic the action of dexamethasone on tyrosine transaminase activity in HTC cells. Surprisingly, in conditions where dexamethasone produced a 10–15-fold increase of tyrosine transaminase activity, neither of the 17 β -carboxamide steroids at concentrations up to 10 μ M significantly modified the baseline activity of this enzyme (shaded bars in Fig. 2*a*). However, when assayed for receptor binding in HTC cell cytosol, the synthetic steroids competed with ³H-labelled dexamethasone as they did in rat liver extracts. The order of affinity was the same. Tyrosine transaminase induction experiments were therefore carried out in which HTC cells were incubated with mixtures of dexamethasone and 17 β -carboxamide steroids. We found that, although devoid of intrinsic glucocorticoid activity, DXH and DMP could inhibit the stimulating effect of dexamethasone on tyrosine transaminase by more than 85% (open bars, Fig. 2*a*). DXB was slightly less inhibitory than DXH because it had a very small agonist activity of its own. The other steroids also antagonised the glucocorticoid effect, in keeping with their affinity for the HTC receptor (Fig. 2*a*). The latter observation suggests that the inhibition of tyrosine transaminase is not a nonspecific action of high concentrations of the dexamethasone derivatives. Data in Fig. 2*b* support this interpretation, as it shows that the anti-glucocorticoid effect is dose dependent and that the order of potency of the inhibitors is consistent with their affinity for the HTC receptor.

Unlike natural corticosteroids, their 9 α -fluoro-derivatives do not bind to transcortin, the plasma corticosteroid-binding globulin. We examined whether this difference was maintained after carrying out the 17 β -substitutions described here. Plasma

Table 1 Natural and synthetic steroids used in this study

Compound*	1-ene	R ₁	R ₂	R ₃	R ₄
Cortisol	—	H	H	OH	CH ₂ OH
DCAC	—	H	H	OH	OH
DCACP	—	H	H	OH	NH-(CH ₂) ₂ -CH ₃
Corticosterone	—	H	H	H	CH ₂ OH
HCAC	—	H	H	H	OH
Dexamethasone	+	F	CH ₃	OH	CH ₂ OH
DXO	+	F	CH ₃	OH	OH
DXM	+	F	CH ₃	OH	NH-CH ₃
DXP	+	F	CH ₃	OH	NH-(CH ₂) ₂ -CH ₃
DXN	+	F	CH ₃	OH	NH-(CH ₂) ₂ -NH-(CH ₂) ₂ -CH ₃
DXH	+	F	CH ₃	OH	NH-(CH ₂) ₅ -CH ₃
DXB	+	F	CH ₃	OH	NH-CH ₂ -C ₆ H ₅
Desoximethasone	+	F	CH ₃	H	CH ₂ OH
DMP	+	F	CH ₃	H	NH-(CH ₂) ₂ -CH ₃
Triamcinolone	+	F	OH	OH	CH ₂ OH

* Trivial names and abbreviations used in text and figures.



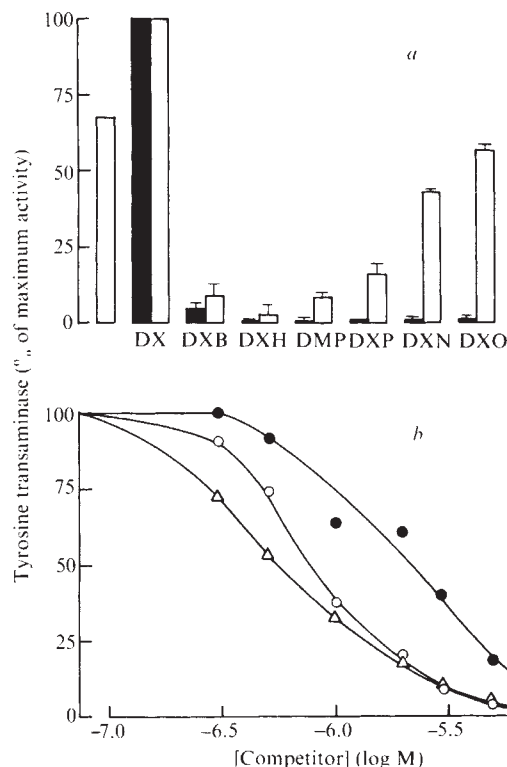


Fig. 2 Anti-glucocorticoid activity of 17 β -carboxamide derivatives of dexamethasone (DX). Rat hepatoma tissue culture (HTC) cells were resuspended in serum-free medium containing 0.1% (w/v) bovine serum albumin and incubated at 37 °C for 16 h in the absence of steroid or in the presence of 10 μ M dexamethasone to determine, as described previously⁴, the maximum tyrosine transaminase activity of the culture. The latter was 11.4–23.7 mU per mg protein, depending on the experiment, whereas it was 1.7–2.3 mU per mg protein in untreated cultures. Parallel cultures were also exposed to the other steroids, alone or in combination (for nomenclature, see Table 1). *a*, Shaded bars: cultures exposed to 10 μ M of a single steroid, as indicated. Open bars: cultures exposed to 25 nM dexamethasone alone (first bar on the left) or in combination with 10 μ M of the steroid indicated. Values given are the range obtained from two to four experiments. *b*, Effect of varying concentrations of the three most potent 17 β -carboxamide derivatives on tyrosine transaminase activity in cultures exposed to 25 nM dexamethasone. Δ , DXB; $\circ$, DXH; $\bullet$, DMP.

from adrenalectomised rats was incubated with ³H-labelled corticosterone and binding to transcortin determined as described elsewhere⁵. We found that a 1,250-fold excess of HCAC, DCAC and DCACP inhibited corticosterone binding by 94, 77 and 71%, respectively, whereas the 9 α -fluoro-derivatives had either no effect or inhibited the binding by only 6–9% (DXB, DXN).

Our results show that glucocorticoid antagonists which bind to the receptor with high affinity can be derived from the synthetic agonists dexamethasone and desoximethasone. Some of the binding properties of these novel steroids are consistent with structure–activity relationships previously defined for the HTC cell glucocorticoid receptor⁷. First, the data confirm the importance of the steroid side chain in assigning both the potency (affinity for the receptor) and the activity class (agonist or antagonist). We now find that the loss of potency on removal of the carbon atom in position 21 can be partially restored by coupling various amines at this position. Crystallographic data (M. Devos, P.F. and P.L., unpublished) make it unlikely that the nitrogen atom could substitute for the oxygen atom of the 21-hydroxyl group as an acceptor in a putative hydrogen bond with the receptor. Rather, binding of the 17 β -carboxamide

steroids would involve the hydrophobic interaction of the new side chain with the receptor site, as the affinity increases with increasing hydrophobicity of the amine. However, other substituents on the steroid skeleton retain their typical influence on the binding: DCACP and DXP have a potency ratio similar to that of the parent compounds cortisol and dexamethasone. Also, the fact that DMP has a greater affinity than DXP shows that the 17 α -hydroxyl group has the same unfavourable effect on the binding to the HTC receptor in 17 β -carboxamide derivatives as in the pregnene series⁷. The persistence of this effect in the absence of a C-21 carbon atom suggests that it could result from a specific influence of the 17 α -hydroxyl group on the orientation of the C₂₀ ketone as suggested earlier⁸. Finally, our results show that conversion of an agonist into an antagonist is compatible with modifications (1–2 unsaturation; 9 α -fluor-, 16 α -methyl- and 11 β -hydroxyl-substitutions) which promote high affinity for the receptor and/or prevent the steroid from binding to plasma transcortin and to cellular proteins other than the receptor. This has important implications for the development of more potent anti-glucocorticoids and receptor-specific ligands possessing chemically reactive functions which would be very useful for receptor purification by affinity chromatography.

G.R. is Maître de Recherches of the FNRS (Belgium) and recipient of grants from the Belgian Foundation for Scientific Medical Research (3.4514.75) and the Ford Foundation (7700312). P.F. and P.L. were supported by grants from the CNRS (LA 04264 to Professor Biserte) and the UER III.

GUY G. ROUSSEAU*
JOSEF KIRCHHOFF†

Université de Louvain
and
International Institute of Cellular
and Molecular Pathology,
ICP 7529, Avenue Hippocrate,
75, B-1200 Brussels, Belgium

PIERRE FORMSTECHE
PATRICK LUSTENBERGER

Laboratoire de Biochimie Structurale
Faculté de Médecine de Lille,
1 Place de Verdun,
59045 Lille Cedex, France

Received 15 January; accepted 12 March 1979.

* To whom correspondence should be addressed.

† Present address: Institut für Physiologische Chemie, D-4300 Essen, 1, FRG.

1. Baxter, J. D. *Pharmac. Ther.* **B2**, 605–659 (1976).
2. Thompson, E. B., Tomkins, G. M. & Curran, J. F. *Proc. natn. Acad. Sci. U.S.A.* **56**, 296–303 (1966).
3. Rousseau, G. G. *J. Steroid Biochem.* **6**, 75–89 (1975).
4. Samuels, H. H. & Tomkins, G. M. *J. molec. Biol.* **52**, 57–74 (1970).
5. Rousseau, G. G., Baxter, J. D. & Tomkins, G. M. *J. molec. Biol.* **67**, 99–115 (1972).
6. Raynaud, J. P. *et al. J. Steroid Biochem.* **6**, 615–622 (1975).
7. Rousseau, G. G. & Schmit, J. P. *J. Steroid Biochem.* **8**, 911–919 (1977).
8. Schmit, J. P. & Rousseau, G. G. *J. Steroid Biochem.* **9**, 909–920 (1978).

Guinea pig prostate is a rich source of nerve growth factor

NERVE GROWTH FACTOR (NGF) is essential for the development of sympathetic nerve cells¹. The richest known source of NGF is the submaxillary gland of the adult male mouse², and NGF from this source has been completely purified and thoroughly characterised³. The active entity (β -subunit) is a protein comprising a dimer of identical 118-residue chains^{4,5}. Another form of mouse NGF, the 2.5S form, has been isolated⁶ and shown to differ from the β entity in that some of the

monomeric chains have undergone enzymatic cleavage of the amino-terminal octapeptide and carboxy-terminal arginine residues^{3,4,7-9}. The 2.5S dimer therefore contains both 118-residue chains and shortened chains. The salivary gland of the mouse seems to be unique in that the corresponding glands of the rat, guinea pig, cow, pig, rabbit and man contain no NGF^{10,11}. No other clearly characterised source of NGF in mammalian tissue *in vivo* has been reported. We now report a new, rich source of NGF—the prostate glands of the adult male guinea pig.

NGF activity in soluble extracts of guinea pig prostate glands (dorsal and ventral lobes) was measured using a biological assay¹²⁻¹⁴ in which the NGF content of a tissue is assessed by the ability of an extract to promote the outgrowth of nerve fibres from explants of dorsal root ganglia from 8-d chick embryos cultured for 24 h in specified conditions. The assay gives the NGF concentration in terms of a biological unit (BU), defined as the amount of NGF required to produce optimum fibre outgrowth. Assuming the same specific activity for guinea pig prostate and mouse submaxillary NGFs, the guinea pig gland is found to contain 1.8 ± 0.5 mg NGF per g wet weight (mean $\pm$ s.e.m., range 340–9,925 μ g, $n = 22$). This organ therefore seems to be of similar potency to the mouse submaxillary gland. The levels in the latter organ (0.3 to 5 mg NGF per g wet weight) are known to vary according to the strain of mouse¹⁵ and the extraction procedure.

Both whole antiserum and antibodies purified by affinity chromatography¹⁶ against mouse submaxillary gland NGF were able to inhibit completely the biological activity of the prostate extracts in tissue culture (see ref. 17). The amounts of pure antibody required to block 1 BU ml⁻¹ *in vitro* of purified mouse NGF and crude extracts of prostate glands were not significantly different (0.62–1.25 μ g ml⁻¹ and 0.31–0.62 μ g ml⁻¹).

Extracts of guinea pig prostate gland were tested by immunodiffusion against antiserum to purified mouse NGF (Fig. 1), and thereby compared with mouse NGF. The interaction of the precipitin lines shows that there are antibodies that recognise determinants common to both mouse and guinea pig NGFs, but that the mouse antigen has determinants not present in the guinea pig NGF (spur formation). This implies that a substantial proportion (note length and intensity of spurs) of the antibodies recognise mouse but not guinea pig NGF. The common antigenic sites may be related to the biologically effective site(s) of the proteins, the structure of which might be expected to be conserved, while the other part(s) of the molecule may be more variable between species.

The NGFs were also compared by immunoprecipitation with purified antibodies against mouse NGF, followed by SDS-gel electrophoresis (Fig. 2). Mouse submaxillary glands and guinea pig prostate glands were extracted in the presence of enzyme inhibitors (see Fig. 2 legend) to prevent the enzymatic cleavage of NGF in the crude extracts. In these conditions, most of the mouse NGF remains as the β -subunit (molecular weight of the monomer 13,259, refs 3–5), although a significant proportion exists in the 2.5S form in spite of the presence of enzyme inhibitors. (The SDS gel shows both the 118-residue monomers from the β and the 2.5S forms and the shorter monomers from the 2.5S NGF alone.) The greater intensity of the higher molecular weight band indicates that the β form predominates. In contrast, the guinea pig NGF shows a single band in the same position as that of the 118-residue monomer from mouse NGF. Omission of enzyme inhibitors during homogenisation gives rise to an increase in the proportion of the 2.5S form of mouse NGF while the guinea pig material remains as the β -subunit. Of several possible explanations, two obvious ones are either that the guinea pig factor is less susceptible to enzymatic cleavage or that the appropriate proteases are present at lower levels in the prostate gland.

The levels of NGF in the guinea pig prostate gland have also been estimated using a two-site radioimmunoassay based on purified antibodies against mouse NGF¹⁸. The radioimmunoassay and biological assay results were in reasonable agreement,

around 35% of the levels expected from the biological assay being detected. The major cause of the discrepancy seems to be the use of antibodies raised against mouse NGF, only some of which recognise guinea pig NGF (see immunodiffusion results), and the capacity of the radioimmunoassay system is thereby substantially reduced. Details of the use of the radioimmunoassay to measure NGF from species other than the mouse will be published elsewhere.

The observation that only a certain proportion of the antibodies to mouse NGF recognise guinea pig NGF contrasts with the similarity of the amount of antiserum required to block 1 BU of guinea pig NGF *in vitro* to the homologous titre. This could be due either to differences in the specific activities of the factors from the two sources or, possibly, to differences in the relationship between the antigenic sites and the sites associated with biological activity.

The significance of the presence in the accessory sex organs of agents with such potent effects on the sympathetic nervous system is not clear, but it is possible that the NGFs have a physiological role unconnected with, and in addition to, their effects on neuronal tissues. The levels of NGF in the mouse submaxillary gland show a well characterised developmental increase² and a marked sex difference¹⁹⁻²¹. A similar situation seems to apply in the reproductive system of the guinea pig. The high levels of NGF *in vivo* thus appear after the development of the sympathetic nervous system^{22,23} and in the absence of any significant sexual dimorphism in the responsive neuronal tissues^{1,24,25}. The actions of NGF on the sympathetic nervous system may require only low levels of the agents, which are attained (although not necessarily detected with the assays

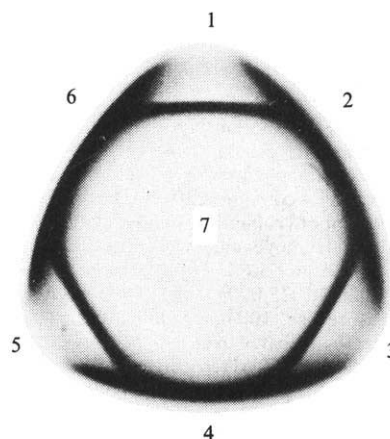


Fig. 1 Comparison of guinea pig and mouse NGFs by immunodiffusion. Wells 1,3,5: 20 μ l of an extract of guinea pig prostate gland, containing 240 μ g NGF per ml (by biological assay). The tissue extract was prepared in distilled water and the lyophilised material was further extracted in 50 mM ammonium acetate/acetic acid buffer pH 4.0, containing 1 mg ml⁻¹ bovine serum albumin, 200 μ g ml⁻¹ sodium azide, 50 μ g ml⁻¹ pancreatic trypsin inhibitor and 5 mM phenylmethanesulphonyl fluoride. The solution was centrifuged (22,000g, 30 min) to remove insoluble material, returned to pH 7.2 with NaOH, centrifuged again to clarify, and stored at -70°C . This procedure removed many proteins from the crude extract but all the NGF was recovered, as judged by biological and radioimmunological assays. Wells 2,4,6: 20 μ l of a solution (in 50 mM Tris-HCl buffer pH 8.5) of purified mouse NGF¹⁸, at a concentration of 250 μ g NGF per ml. Well 7: 20 μ l of a solution (in 50 mM Tris-HCl buffer pH 8.5) of lyophilised sheep antiserum to mouse NGF¹⁸ (60 mg ml⁻¹). Immunodiffusion plates were made up of 0.4% agarose containing 0.1% Thimerosal. Plates were incubated at 20°C in a humidified atmosphere for 48 h, washed for 5 d in 150 mM NaCl solution to remove non-precipitated proteins, stained for 1–2 h in 0.25% Coomassie brilliant blue in 150 mM saline, destained for 3 d in 150 mM saline, and photographed. An identical result was obtained when wells 2, 4 and 6 contained an extract, prepared as described above, of mouse submaxillary gland, containing 160 μ g NGF per ml.

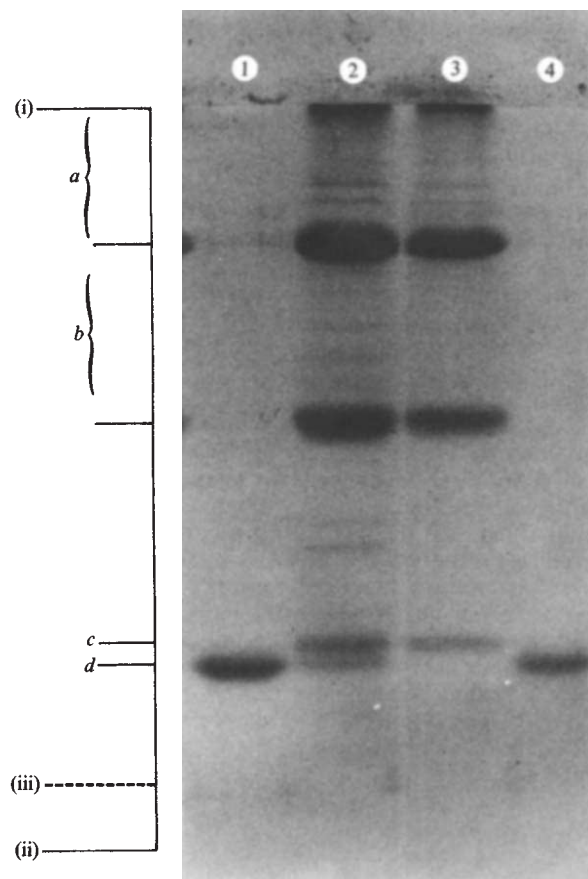


Fig. 2 SDS-gel electrophoresis of immunoprecipitates of guinea pig and mouse NGFs. Two guinea pig prostate glands (7 ml) and two mouse submaxillary glands (5 ml) were homogenised with a Potter ground glass homogeniser in the volumes shown of 50 mM Tris-HCl buffer pH 7.4 containing 1% Triton X-100, 300 $\mu\text{g ml}^{-1}$ phenylmethanesulphonyl fluoride, and 20 $\mu\text{g ml}^{-1}$ pancreatic trypsin inhibitor. Homogenates were centrifuged at 22,000g (30 min) and stored at -50°C until required; the supernatants were recentrifuged at 25,000g for 20 min after thawing and immediately before use. 100 μl of guinea pig extract and 60 μl mouse extract were incubated overnight at 4°C with 10 μl Nonidet P40 detergent, 250 and 230 μl respectively of a solution (in 50 mM Tris-HCl buffer pH 7.2, containing 150 mM NaCl) of purified sheep antibodies¹⁶ to mouse 2.5S NGF (0.73 mg antibodies per ml) and Tris-NaCl buffer to a final volume of 400 μl . Pilot experiments determined that these ratios of antigen to antibody were at the equivalence point. Immunoprecipitates were collected by centrifugation at 15,000g for 10 min and washed three times with 500 μl Tris-NaCl buffer containing 2.5% (v/v) Nonidet P40. The immunoprecipitates were then dissolved in the sample buffer for electrophoresis (15% glycerol, 5% 2-mercaptoethanol, 2.25% SDS, 0.01% bromophenol blue and 0.0625 M Tris-HCl buffer pH 6.8) by heating at 90°C for 10 min. The solubilised immunoprecipitates were subjected to SDS-15% polyacrylamide gel electrophoresis according to the method of Laemmli²⁹. The slab gel was performed in an apparatus similar to that described by Studier³⁰. The slab gel was stained with 0.25% Coomassie brilliant blue in methanol acetic acid water (5:1:5 by volume) solution and destained in the same solution. (i), (ii), Top and bottom respectively of the SDS-polyacrylamide gel; (iii), bromophenol blue front; a, immunoglobulin heavy chains; b, immunoglobulin light chains; c, position of the intact monomer of mouse β and 2.5S NGF, and of the guinea pig NGF; d, position of the smaller monomer from mouse 2.5S NGF. 1, Purified mouse NGF, 4 μg . Both chains are entirely converted to the smaller chain, as found in 2.5S NGF, when NGF is purified according to ref. 18. 2, Mouse NGF immunoprecipitate. 3, Guinea pig NGF immunoprecipitate. 4, Purified mouse NGF 2 μg (as above).

presently available) in developing (and female) animals. Another, possibly sex-linked, function is conceivable for these proteins, for which very high levels of NGF are required.

From a practical point of view, the high levels of NGF in the guinea pig prostate gland make this source an attractive alternative to the male mouse submaxillary glands. Interestingly, in view of the appearance of an epidermal growth factor (EGF) in the mouse submaxillary gland together with the NGF²⁶⁻²⁸, preliminary studies indicate that the guinea pig prostate gland could be a source of EGF of a potency comparable with or greater than the mouse submaxillary gland.

We thank Professor M. Schachter, Dr Barbara Banks, Dr F. L. Pearce and Professor H. Thoenen for valuable discussions. Financial assistance was given to G.P.H. by the SRC, the Wellcome Trust, and EMBO.

G. P. HARPER*†
Y. A. BARDE†
G. BURNSTOCK‡
J. R. CARSTAIRS*
M. E. DENNISON‡
K. SUDA†
C. A. VERNON*§

* Department of Chemistry,
University College London,
20 Gordon Street, London WC1, UK

† Department of Pharmacology,
Biocentre of the University,
Klingelbergstrasse 70,
CH-4056 Basel, Switzerland

‡ Department of Anatomy and Embryology,
University College London,
Gower Street, London WC1, UK

Received 27 October 1978; accepted 27 March 1979.

§ To whom reprint requests should be addressed.

1. Levi-Montalcini, R. & Angeletti, P. U. *Physiol. Rev.* **48**, 534-569 (1968).
2. Cohen, S. *Proc. natn. Acad. Sci. U.S.A.* **46**, 302-311 (1960).
3. Server, A. C. & Shooter, E. M. *Adv. Protein Chem.* **31**, 339-409 (1977).
4. Angeletti, R. H. & Bradshaw R. A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2417-2420 (1971).
5. Greene, L. A. Varon, S., Piltch, A. & Shooter, E. M. *Neurobiology* **1**, 37-48 (1971).
6. Bocchini, V. & Angeletti, P. U. *Proc. natn. Acad. Sci. U.S.A.* **64**, 787-794 (1969).
7. Moore, J. B., Mobley, W. C. & Shooter E. M. *Biochemistry* **13**, 833-840 (1974).
8. Mobley, W. C., Schenker, A. & Shooter, E. M. *Biochemistry* **15**, 5543-5552 (1976).
9. Angeletti, R. H., Bradshaw, R. A. & Marshall, G. R. *Int. J. Peptide Protein Res.* **6**, 321-328 (1974).
10. Banks, B. E. C. *et al. Nature* **246**, 503-504 (1973).
11. Harper, G. P., Pearce, F. L. & Vernon, C. A. (in preparation).
12. Varon, S., Nomura, J., Perez-Polo, J. R. & Shooter, E. M. *Meih. Neurochem.* **3**, 203-229 (1972).
13. Fenton, E. L. *Expl. Cell. Res.* **59**, 383-392 (1970).
14. Pearce, F. L., Banthorpe, D. V., Cook, J. M. & Vernon, C. A. *Eur. J. Biochem.* **32**, 569-575 (1973).
15. Bamberg, J. R., Derby, M. A. & Shooter, E. M. *Neurobiology* **1**, 115-120 (1971).
16. Stoeckel, K., Gagnon, C., Guroff, G. & Thoenen, H. *J. Neurochem.* **26**, 1207-1211 (1976).
17. Bailey, G. S. *et al. Biochim. biophys. Acta* **437**, 259-263 (1976).
18. Suda, K., Barde, Y. A. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4042-4046 (1978).
19. Caramia, F., Angeletti, P. U. & Levi-Montalcini, R. *Endocrinology* **70**, 915-922 (1962).
20. Bueker, E. D., Weis, P. & Schenkein, I. *Proc. Soc. exp. Biol. Med.* **118**, 204-207 (1965).
21. Goldstein, M. N. & Burdman, J. A. *Anat. Rec.* **151**, 199-208 (1965).
22. de Champlain, J., Malmfors, T., Olson, L. & Sachs, C. *Acta physiol. scand.* **80**, 276-288 (1970).
23. Owman, C., Sjöberg, N. O. & Swedin, G. *Zellforsch.* **116**, 319-341 (1971).
24. Kato, D. *Yonaga acta medica* **16**, 172-175 (1973).
25. Levi-Montalcini, R. *Harvey Lect.* **60**, 217-259 (1966).
26. Cohen, S. *Natn. Cancer Inst. Monogr.* **13**, 13-37 (1964).
27. Cohen, S. *Dev. Biol.* **12**, 394-407 (1965).
28. Cohen, S. & Taylor, J. M. *Macromolecules Regulating Growth and Development*, 25-42 (Academic, New York, 1974).
29. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
30. Studier, F. W. *J. molec. Biol.* **79**, 237-248 (1973).

Lipidic intramembranous particles

THE nature of intramembranous particles has recently been discussed in relation to the mode of freeze fracturing^{1,2}. It has been argued that particles that do not show complementarity (complementary pits on the opposite fracture face) are a reflection of protein penetrating the membrane. The protein

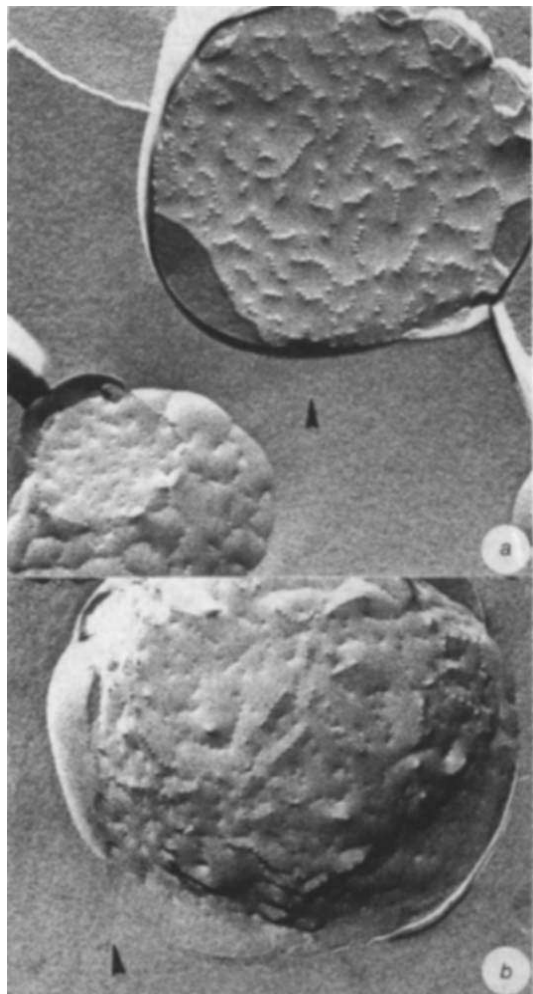


Fig. 1 An equimolar mixture of egg-lecithin, isolated according to established procedures, and cardiolipin (bovine heart cardiolipin) with a Ca^{2+} /cardiolipin molar ratio of 0:1 is dissolved in ether and injected according to the method described by Deamer and Bangham⁸ in a buffer containing 150 mM NaCl and 25 mM Tris, pH 7.5. The purity of the lipids was found to be at least 99% as evidenced by TLC on silica-gel G using the elution system chloroform/methanol/acidic acid (93:30:12). Shadowing direction is indicated by arrows. $\times 52,000$.

giving rise to an intramembraneous particle is plastically deformed during the fracturing procedure and is probably pulled out to one of the two fracture faces. Moreover we have argued that particles showing complementarity (pits) are possibly of lipidic origin. This latter hypothesis was mainly based on the determination of the nature of the intramembraneous particles of the outer membrane of *Escherichia coli*. These complementary particles were shown to be determined by lipopolysaccharide^{3,4}. We show here that lipid can by itself form intramembraneous particles showing complementarity and that these particles may be inverted micelles of phospholipid sandwiched between lipid monolayers.

To test the hypothesis that lipid can give rise to intramembraneous particles with complementary impressions on the opposite fracture face, we have studied primarily lipids that can exist in non-bilayer phases. Good candidates are phosphatidylethanolamines from natural sources⁵ and especially cardiolipin in the presence of Ca^{2+} (ref. 6) which can form the hexagonal phase II.

Recently we have shown that the transition of cardiolipin from the bilayer configuration to the hexagonal phase by addition of Ca^{2+} occurs by an intermediate phase of particles about 70 Å in diameter, as visualised by freeze fracturing⁷. This intermediate phase is characterised by the ³¹P NMR spectrum

which indicates isotropic motional averaging. This phase, which exists at Ca^{2+} -cardiolipin ratios below 0.6, can be interpreted as inverted micelles, that is particles with hydrophilic cores and hydrophobic edges.

To see whether such inverted micelles can form intramembraneous particles we have prepared model membranes of an equimolar mixture of lecithin and cardiolipin at a low Ca^{2+} cardiolipin ratio as described in Fig. 1 legend.

Freeze fracturing shows fracture faces of bilayers organised predominantly in large unilamellar structures. Also some multilamellar structures are found. On the fracture faces of these bilayers particles of a rather uniform size are present. The mean diameter is 100 Å (s.d. 14 Å; $N = 100$). The pits are about 30 Å smaller. Fracture faces bearing exclusively particles or exclusively pits are found (Fig. 1a) but also fracture faces with both particles and pits can be observed (Fig. 1b). If Ca^{2+} is removed by EGTA extraction, neither particles nor pits are observed—only smooth fracture faces. On the other hand excess Ca^{2+} results in a homogeneous conglomerate of globules about 100 Å in diameter. (These latter results will be published in detail elsewhere.)

The most important conclusion that can be drawn from these observations is that lipid by itself can form intramembraneous particles and that such particles show complementarity. These particles might arise from invaginations of the bilayer, but this is unlikely as one should expect a variation in the size of particles and consequently in the size of the pits, which is not the case. In our opinion these particles represent inverted micelles of phospholipid sandwiched between lipid monolayers.

A. J. VERKLEIJ

Institute of Molecular Biology,

C. MOMBERS

Biochemical Laboratory,

J. LEUNISSEN-BIJVELT

P. H. J. Th. VERVERGAERT

Department of Molecular Cell Biology,

Section Electron Microscopy,

State University of Utrecht, Transitorium 3,

Padualaan 8 Utrecht, The Netherlands

Received 14 November 1978; accepted 16 March 1979.

1. Vergeraert, P. H. J. T. & Verkleij, A. J. *Experientia* **34**, 354–355 (1978).
2. Verkleij, A. J. & Vergeraert, P. H. J. T. *Biochim. biophys. Acta* **515**, 303–327 (1978).
3. Verkleij, A. J., van Alphen, L., Bijvelt, J. & Lugtenberg, B. *Biochim. biophys. Acta* **426**, 581–586 (1976).
4. van Alphen, L., Verkleij, A. J., Leunissen-Bijvelt, J. & Lugtenberg, B. *J. Bact.* **134**, 1089–1098 (1978).
5. Reiss-Husson, F. *J. molec. Biol.* **25**, 363–374 (1967).
6. Rand, R. P. & Sengupta, S. *Biochim. biophys. Acta* **255**, 484–492 (1972).
7. Cullis, P. R., Verkleij, A. J. & Vergeraert, P. H. J. T. *Biochim. biophys. Acta* **513**, 11–20 (1978).
8. Deamer, D. W. & Bangham, A. D. *Biochim. biophys. Acta* **433**, 629–634 (1976).

The role of spectrin in erythrocyte membrane-stimulated actin polymerisation

PROTEINS on the cytoplasmic surface of the erythrocyte membrane, including spectrin and actin, are postulated to comprise the red cell cytoskeleton^{1–3}, but little is known about the role of actin or its association with the membrane. We have reported that monomeric (G) actin added to erythrocyte ghosts selectively associates with a component at the cytoplasmic surface of the membrane^{4,5}. We now show that this component is unlikely to be spectrin and that actin binding occurs by stimulated actin polymerisation.

Large amounts of added G-actin bind to the cytoplasmic surface of unsealed ghosts, but little or none binds to sealed ghosts or spectrin-actin-depleted inside-out vesicles^{4,5}. After spectrin-actin-depleted inside-out vesicles (referred to here

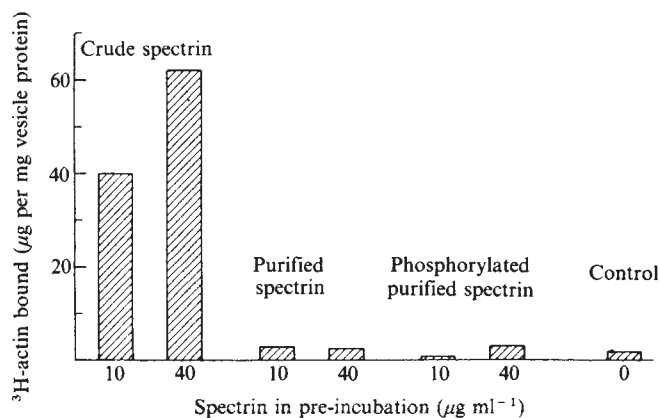


Fig. 1 Actin binding to vesicles reconstituted by preincubation with spectrin. Unsealed, spectrin-actin-depleted inside-out vesicles were prepared from ghosts¹⁸ of freshly drawn human erythrocytes by one wash in ice-cold 0.3 mM sodium phosphate buffer, pH 7.6, incubation for 40 min in 40 volumes of 0.3 mM sodium phosphate buffer at 37 °C and two washes in ice-cold 0.3 mM sodium phosphate buffer. Crude spectrin (primarily spectrin, actin and band 4.1) was prepared by extracting ghosts in 0.3 mM sodium phosphate buffer, 37 °C and phosphorylated by incubating 1 mg spectrin in 1 ml of 100 mM NaCl, 10 mM Tris pH 7.4, 5 mM MgCl₂, 50 µM ATP containing 1 µCi of ³²P-ATP and 200 µg erythrocyte kinase¹⁹ at 37 °C for 30 min. Saturation of ³²P incorporation occurred after ~20 min. (33 mol PO₄ per mol spectrin in this experiment.) After phosphorylation, the ³²P-spectrin was purified on a sucrose density gradient⁶ (similar results were obtained using spectrin which was phosphorylated after gradient purification). Vesicles (0.3 mg protein per ml) were incubated for 1 h on ice with the indicated concentrations of spectrin in 20 mM KCl, 2 mM MgCl₂, 5 mM sodium phosphate pH 7.6, and 0.75 mM β-mercaptoethanol, washed once and resuspended in this same medium to a protein concentration of 0.25 mg ml⁻¹. Control experiments confirmed that the amounts of crude, pure or phosphorylated pure spectrin which bound to vesicles were comparable when they were added at the same concentration (~60 µg per mg vesicle protein and 90 µg per mg for added spectrin concentrations of 10 and 40 µg ml⁻¹ respectively). Preparation of [³H]G-actin from rabbit skeletal muscle and quantification of ³H-actin binding as described previously⁵. Briefly, 40 µg ml⁻¹ [³H]G-actin was incubated in a volume of 0.8 ml with 0.25 mg ml⁻¹ membrane protein in 20 mM KCl, 5 mM sodium phosphate pH 6.5, 2 mM MgCl₂, 1 mM ATP and 0.75 mM β-mercaptoethanol (actin-binding buffer) on ice for 1 h. Samples were layered on 0.2 ml of 20% w/v sucrose (made in actin-binding buffer) in 0.5 ml polyethylene microfuge tubes, centrifuged at 18,000 r.p.m. for 25 min in a Sorvall SS34 rotor and the tips, containing the membrane pellet, frozen in liquid N₂, cut off, placed in 10 ml of scintillation fluid and counted for ³H. No ³H-actin sedimented through the sucrose in the absence of membranes in any of the conditions studied.

simply as vesicles) were reconstituted with crude spectrin (primarily spectrin, actin and band 4.1; for nomenclature see ref. 1), their ability to bind added G-actin became nearly equal to that of unsealed ghosts⁵. As the major component of crude spectrin is spectrin itself, we tested the ability of the purified protein to stimulate actin binding (Fig. 1). Crude spectrin was effective at stimulating actin binding to inside-out vesicles at low concentrations, but purified spectrin had little or no effect, even when added in concentrations sufficient to saturate the spectrin binding capacity (ref. 6, and our unpublished data) of the vesicles.

As the interaction of spectrin and actin has been reported to be sensitive to spectrin's state of phosphorylation⁷, we considered the possibility that the purified spectrin was inactive because it had become dephosphorylated. However, reconstitution of vesicles with phosphorylated spectrin did not stimulate actin binding (Fig. 1).

The ability of crude, but not purified spectrin to stimulate actin binding prompted us to explore the possibility that some component of crude spectrin other than spectrin itself was responsible for this effect. Crude spectrin was centrifuged on a 12-ml linear 5–20% sucrose gradient which was fractionated into 14 aliquots. Samples of vesicles were reconstituted with each of the gradient fractions and tested for their ability to bind added G-actin (Fig. 2). The only fraction capable of stimulating binding was found near the bottom of the gradient and contained spectrin, red-cell actin, and band 4.1. The ability of this fraction to stimulate binding was less than an equivalent amount of crude spectrin, indicating some loss of stimulatory activity during purification.

These results show that some component of crude spectrin other than spectrin itself, can stimulate actin binding to red-cell membranes. We propose that this stimulated binding occurs as a direct result of induced polymerisation of the added G-actin because: (1) actin binding depends on the presence of Mg. (2) Actin binding is non-saturable. (3) Actin binding can be prevented by DNase I and cytochalasin B⁵. (4) Actin binding is accompanied by terminal phosphate hydrolysis from the tightly bound ATP of the actin molecule⁸, a hallmark of actin polymer formation. In addition, we have shown that vesicles reconstituted with crude spectrin and subsequently incubated with G-actin become covered with actin filaments which emanate from their surface⁵. Decoration of these filaments with heavy meromyosin (Fig. 3) reveals that they all have the same polarity, arrow heads pointing towards the membrane, indicating that although polymerisation is nucleated by the membrane, growth proceeds at the free, distal end of the filament⁹. Thus, the polarity of the filaments is opposite to that seen *in situ* in other cell types^{10–13}.

In line with our conclusion that crude spectrin stimulates actin binding to vesicles by polymerisation, we have used a novel microassay for actin polymerisation^{8,14} and found that crude spectrin can stimulate actin polymerisation in solution, whereas purified spectrin, whether phosphorylated or not, does not stimulate actin polymerisation. This suggests that there is a direct relationship between polymerisation-inducing capacity and stimulation of actin binding to membranes.

Our attempt to define the components of crude spectrin responsible for stimulating actin binding requires further exploration, as the gradient fraction capable of stimulating binding contains all three of the major components of crude spectrin (spectrin, red-cell actin and band 4.1). Their presence near the bottom of the gradient suggests that they may be associated in a supramolecular complex migrating with a sedimentation coefficient considerably greater than any of the pro-

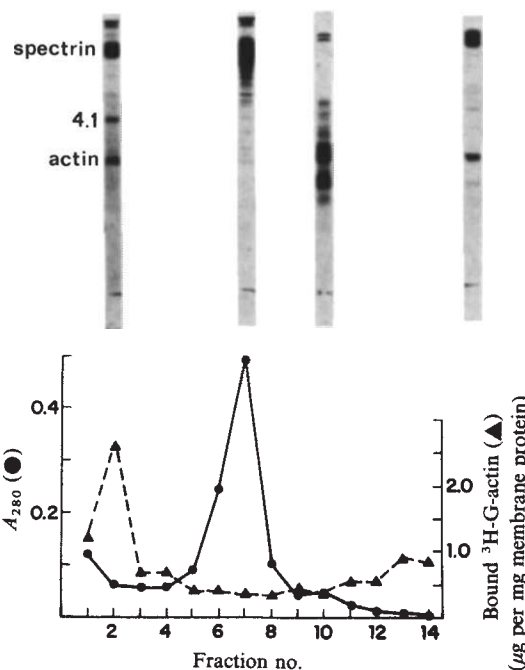


Fig. 2 Stimulatory effect of spectrin gradient fractions on actin binding by vesicles. Crude spectrin was centrifuged on a 5–20% sucrose gradient as described in Fig. 1. The gradient was fractionated, and 200-µl aliquots of each fraction were incubated in 1 ml of 20 mM KCl, 2 mM MgCl₂, 5 mM sodium phosphate pH 7.6, 0.75 mM β-mercaptoethanol containing spectrin-actin-depleted inside-out vesicles (0.3 mg protein) for 1 h at 4 °C. Samples of vesicles reconstituted with each of the gradient fractions were washed once and resuspended in this same buffer to 0.25 mg protein per ml and tested for actin binding as described in Fig. 1. The absorbance of each gradient fraction (solid line) and actin bound by vesicles reconstituted with that fraction (dashed line) are plotted. 5% SDS polyacrylamide gels²⁰ of selected gradient fractions are shown above the corresponding fraction numbers, the gel to the far right being the crude spectrin.



Fig. 3 Electron micrograph of negatively stained inside-out vesicles reconstituted with crude spectrin, incubated with G-actin and subsequently decorated with heavy meromyosin. Spectrin-actin-depleted inside-out vesicles were reconstituted with crude spectrin ($40 \mu\text{g ml}^{-1}$) and incubated with $40 \mu\text{g ml}^{-1}$ G-actin as described in Fig. 1. After incubation for 20 min with the actin, the vesicles were sedimented through 20% w/v sucrose to separate them from unbound actin and applied to a formvar-coated grid. The grids were rinsed three times with 0.6 M KCl, 10 mM Tris pH 7.6 followed by application of 0.2 mg ml^{-1} heavy meromyosin in the same medium. Grids were stained with 1% uranyl acetate. $\times 230,000$.

teins separately, but each of these proteins may be capable of forming multimeric associations with itself (spectrin is known to form tetramers and actin to form filaments). One or both of these factors may contribute to the formation of the active complex, but the presence of actin in the complex is of special significance because small quantities of purified F-actin, when reconstituted onto vesicles by the method used for crude spectrin or the gradient fractions, can stimulate the subsequent binding of added G-actin. Moreover, when purified F-actin is added to G-actin in solution at actin concentrations and solvent conditions identical to those used in our binding studies, rapid and extensive actin polymerisation takes place (data not shown). These findings suggest that crude spectrin and the active gradient fraction may contain F-actin, or actin aggregates, which are responsible for the polymerisation-related binding of added G-actin. This conclusion can be reconciled with the fact that unmodified red-cell ghosts and crude spectrin do not contain structures identifiable as F-actin by proposing that the actin on the membrane, and as extracted in crude spectrin, exists in small aggregates or proto-filaments consisting of three to five monomers. Such F-actin aggregates may be highly efficient in promoting filament formation in the presence of an excess of G-actin¹⁵, both in solution and when bound to red-cell membranes, and would allow red-cell actin to be bound and distributed in a homogeneous and apparently amorphous manner over the membrane surface.

Our hypothesis of membrane-bound actin aggregates can explain the membrane-induced actin polymerisation observed with ghosts and reconstituted inside-out vesicles, as well as crude spectrin-induced actin polymerisation in solution. Our

observation that purified spectrin has none of these activities, however, contradicts the results of Pinder *et al.*⁷, who measured viscosity changes and concluded that phosphorylated spectrin can induce actin polymerisation in solution. To reconcile our observations with those of Pinder *et al.*, we suggest that their spectrin, which was chromatographed on Sephadex G-200, contained red-cell actin aggregates which induced the polymerisation of added G-actin. Observations by other workers^{16,17} provide clear evidence that spectrin prepared from the void volume fractions of this or other large-pore gels will be contaminated by small quantities of actin and other proteins. The effect of spectrin phosphorylation on actin viscosity might be to induce the cross-linking of adjacent filaments which had already been formed.

This research was supported by grants NIH HL 17411 and NSF PCM 77-13962. We thank C. Korsgren and P. Jackson for technical assistance.

CARL M. COHEN*
DANIEL BRANTON

*Cell and Developmental Biology,
The Biological Laboratories,
Harvard University,
Cambridge, Massachusetts 02138*

Received 27 December 1978; accepted 16 March 1979.

* Present address: Department of Research, Division of Hematology, St. Elizabeth's Hospital, 736 Cambridge Street, Boston, Massachusetts 02135.

1. Yu, J., Fischman, D. A. & Steck, T. L. *J. supramolec. Struct.* **1**, 233-248 (1973).
2. Lux, S. E., John, K. M. & Karnovsky, M. J. *J. clin. Invest.* **58**, 955-963 (1976).
3. Sheetz, M. P. & Sawyer, D. J. *J. supramolec. Struct.* **8**, 399-412 (1978).
4. Cohen, C. M., Jackson, P. L. & Branton, D. *J. supramolec. Struct. Suppl.* **2**, 208 (1978).
5. Cohen, C. M., Jackson, P. L. & Branton, D. *J. supramolec. Struct.* **9**, 113-124 (1978).
6. Bennett, V. & Branton D. *J. biol. Chem.* **252**, 2753-2763 (1977).
7. Pinder, J. C., Bray, D. & Gratzner, W. *Nature* **270**, 752-754 (1977).
8. Cohen, C. M. & Branton, D. *J. Cell Biol.* **79**, 262a (1978).
9. Pollard, T. D. in *Molecules and Cell Movement* (eds Inoue, S. & Stephens, R. E.) 259-286 (Raven, New York, 1975).
10. Mooseker, M. S. & Tilney, L. G. *J. Cell Biol.* **67**, 725-743 (1975).
11. Begg, D. A., Rodewald, R. & Rebhun, L. I. *J. Cell Biol.* **79**, 846-852 (1978).
12. Edds, K. T. *Exptl Cell Res.* **108**, 437-456 (1977).
13. Nachmias, V. T. & Ash, A. in *Cell Motility* Vol. 3, (eds Goldman, R., Pollard, T. & Rosenbaum, J.) 771-783 (Cold Spring Harbor, New York, 1976).
14. Cohen, C. M., Atterbury, B. & Branton, D. (in preparation).
15. Oosawa, F. & Kasai, M. in *Subunits in Biological Systems* (eds Timasheff, S. M. & Fasman, G. D.) 261-322 (Marcel Dekker, New York, 1971).
16. Ungewickell, E. & Gratzner, W. *Eur. J. Biochem.* **88**, 379-385 (1978).
17. Dunbar, J. C. & Ralston, G. B. *Biochim. biophys. Acta* **510**, 283-291 (1978).
18. Steck, T. L. & Kant, J. A. *Meth. Enzym.* **31**, 172-180 (1974).
19. Hosey, M. M. & Tao, M. *Biochim. biophys. Acta* **482**, 348-357 (1977).
20. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. *Biochemistry* **10**, 2606-2616 (1971).

Comparison of the predicted model of α -lytic protease with the X-ray structure

THERE are several methods for the prediction of secondary structural features of protein molecules from the amino acid sequence¹⁻⁵. Application of these methods to proteins of unknown tertiary structure has met with mixed results^{6,7}, and it has been suggested that it is the protein structural type which will determine which predictive method will succeed⁸. There is also great interest in the *ab initio* prediction of tertiary structures of proteins⁹⁻¹¹, but this has not yet been achieved. One possible method for predicting the tertiary structure of an enzyme is by the use of a known amino acid sequence and a previously determined tertiary structure of another related homologous isofunctional enzyme. This method was used by McLachlan and Shotton¹² in an attempt to fit the α -lytic protease sequence into the polypeptide chain folding of elastase¹³ and chymotrypsin¹⁴. It has also been used to predict the tertiary structure of tropomyosin C (ref. 15). Now that the structure of α -lytic protease is known¹⁶ to a resolution of 2.8 Å, we can assess the accuracy of

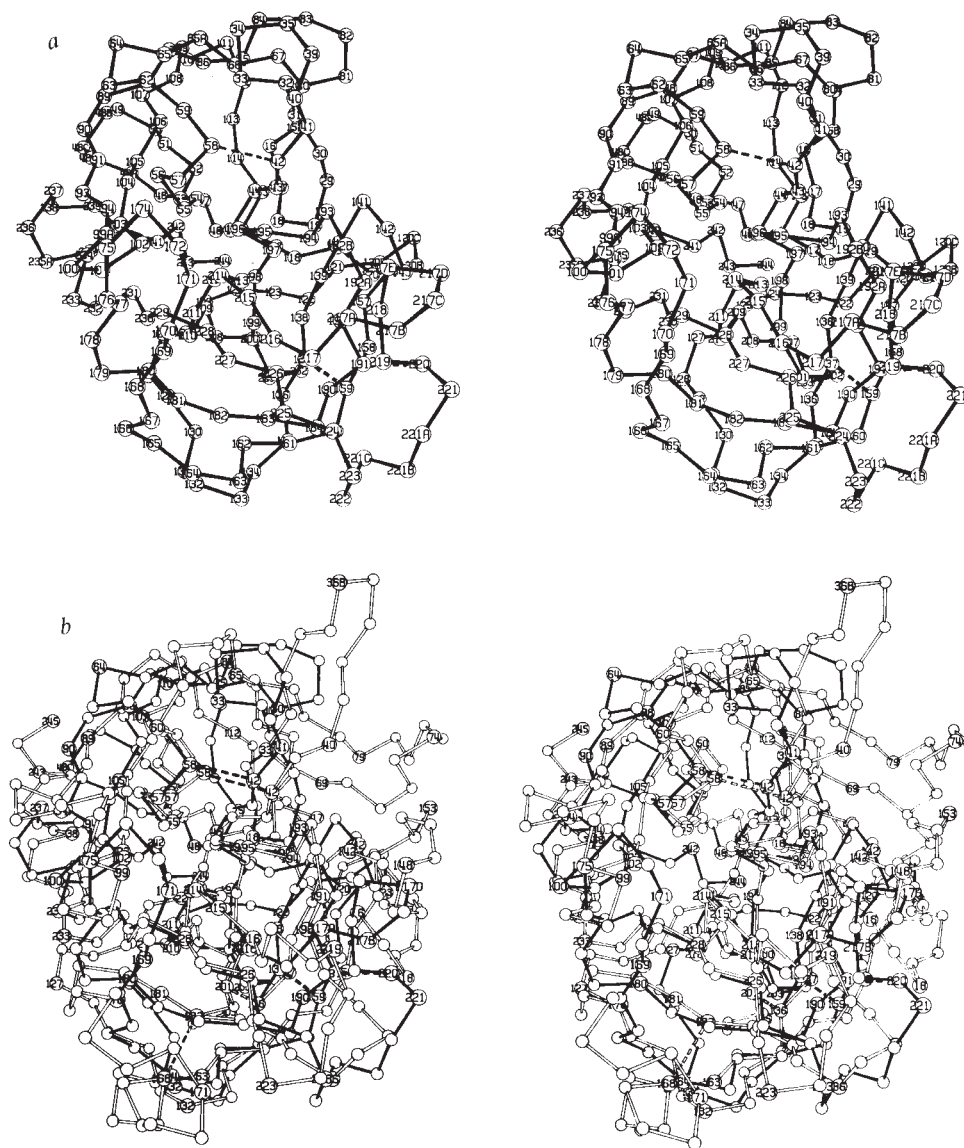


Fig. 1 *a*, A stereo drawing of the α -carbon atom backbone of α -lytic protease, with each α -carbon atom position numbered according to that of chymotrypsinogen A as described in detail in ref. 20. The active site region is located in the central portion of this drawing, where α -carbon atom positions of the catalytic quartet of Ser 214, Asp 102, His 57 and Ser 192 are evident. The three disulphide bridges 42–58, 137–159 and 191–220 present in this molecule are denoted by dashed virtual bonds. *b*, The α -carbon atom drawing of α -lytic protease (solid virtual bonds) is superimposed on that of elastase (open virtual bonds) in a manner designed to maximise topological equivalences²⁴. Every fifth α -carbon atom position of each enzyme is numbered; additional numbering is also present for residues discussed in the text. See Fig. 1 of ref. 12 for more details of the model proposed.

such predictions. The tertiary structures of two related bacterial serine proteases, SGPA and SGPB and their structural relationship to the pancreatic enzymes have been published previously^{17–19}. Many of the conclusions drawn from those studies are also applicable to the present discussion on α -lytic protease and need not be repeated. Rather, we shall consider the basic premise of sequence homology in phylogenetically distant proteins being used to deduce tertiary structures. A recent realignment²⁰ of the amino acid sequences of Gly-Asp-Ser-Gly-Gly proteases, which was based on the known topological equivalences of α -carbon atoms, indicates an overall sequence identity of only 18% between α -lytic protease and elastase. We show here that this low value is independent of the environment of the topologically equivalent polypeptide chains (whether the residues are internal or external) and that the sequence identity is associated with only those residues which are in the immediate vicinity of the active site quartet, Asp 102, His 57, Ser 195 and Ser 214.

The complete amino acid sequence determination of this bacterial serine protease from *Myxobacter* 495 revealed short stretches of sequence which were highly homologous to presumably corresponding regions in the sequences of the mammalian pancreatic enzymes—chymotrypsin, trypsin and elastase. This homology and the similar catalytic behaviour towards peptide substrates²² suggested that α -lytic protease might have a tertiary structure similar to that of the mammalian enzymes despite the fact that it has a much smaller molecular

weight than porcine elastase (19,900 compared to 25,900). In regions of the polypeptide chain of α -lytic protease which do not contain residues of the active site, the sequence homology with the pancreatic enzymes is so low that serious sequence misalignments have resulted^{12,21} and persisted even with the knowledge of the tertiary structure of SGPB¹⁷. In spite of the lack of sequence homology, there is considerable tertiary structural equivalence which is far more extensive and presumably more conserved than the amino acid sequence identities would suggest.

The polypeptide chain folding of α -lytic protease (represented by α -carbon atom positions only) is illustrated in Fig 1*a*. Figure 1*b* shows the α -carbon atoms of α -lytic protease superimposed on the α -carbon atom backbone drawing of elastase²³ (coordinates for elastase are from the Brookhaven Protein Data Bank). The superimposition of these two enzyme structures was achieved using a computer program written by W. Bennett of Yale University and is based on the methods described by Rossmann and Argos²⁴. Examination of Fig. 1*b* shows that there is indeed topological equivalence of α -carbon atoms between α -lytic protease and elastase. However, this tertiary structural homology²⁰ (55% of the α -carbon atoms of α -lytic protease (108 residues) have a topologically equivalent α -carbon atom in elastase within an r.m.s. deviation of 2.08 Å) is limited to the central core, and in particular to those stretches of polypeptide chain which support or contain residues directly involved with substrate binding or the catalytic event.

An alternative representation of the topological equivalence between α -lytic protease and elastase is given in Table 1 which summarises those 76 residues in both domains of these enzymes (domain I, residues 16–120; domain II, 121–245) which have the highest degree of topological equivalence. It is clear from Table 1 that the number of topologically equivalent residues in domain II is more than double that in domain I. This is also true if the topological comparisons are done treating each domain independently. There are only three strands of polypeptide chain which superimpose well in domain I and these strands support the two catalytically important residues His 57 and Asp 102. The apparently high level of amino acid identity (36%) is the result of the small absolute number of topologically equivalent residues (22) in domain I.

Domain II has 54 topologically equivalent α -carbon atoms in six strands of antiparallel β -pleated sheet. The degree of sequence identity between α -lytic protease and elastase is relatively low (22%). There is only one strand (that supporting Ser 195) which has extensive sequence identity, as shown in Table 1. The lack of sequence homology between the first three topo-

logically equivalent strands of domain II is not sufficient to predict a correct model successfully. Indeed, there is only one residue that is structurally required out of the 25 residues comprising these three strands. Gly 140 is structurally invariant²⁰ and is required as such; any side chain at this position would sterically interfere with the side chain of Asp 194, a residue essential to determining the proper catalytically active conformation of serine protease binding sites^{25,26}. The three strands of topologically equivalent residues in domain I of α -lytic protease were correctly predicted from the elastase fold¹² as were strands 4 and 6 domain II. However, the remaining four strands, although topologically equivalent in the two enzymes, had so little sequence homology that the correct register was not discerned. Clearly, what is conserved throughout the evolution of these important digestive molecules is the three-dimensional structure and not the amino acid sequence, except in the immediate vicinity of the catalytic quartet.

The α -lytic protease and pancreatic elastase have a similar substrate specificity^{22,27}. The steric interactions whereby these two structurally related enzymes achieve this specificity for small

Table 1 Topological equivalence and sequence homology of α -lytic protease and elastase

[A] Domain I (residues 16–120)

							41	42	43	44	
α -LP							L	C	S	V	Strand 1
Elastase							T	C	G	G	
Δ							0.93	1.37	1.21	0.94	
		51	52	53	54	55	56	57	58	59	
α -LP		G	F	V	T	A	G	H	C	G	Strand 2
Elastase		W	V	M	T	A	A	H	C	V	
Δ		1.03	0.96	0.49	0.39	0.33	0.46	1.04	1.54	0.58	
	109	108	107	106	105	104	103	102	101		
α -LP	T	L	S	V	W	A	R	D	N		Strand 3
Elastase	A	L	R	L	L	A	I	D	Y		
Δ	1.72	0.89	1.21	0.81	0.43	0.83	1.21	0.57	2.17		

[B] Domain II (residues 121–245)

	131	132	133	134	135	136	137	138	139	140	141	142		
α -LP	A	V	G	A	A	V	C	R	S	G	R	T	Strand 1	
Elastase	A	N	N	S	P	C	Y	I	T	G	W	G		
Δ	1.32	1.87	1.57	1.53	1.69	0.50	1.20	0.95	1.11	0.70	1.54	1.82		
			163	162	161	160	159	158	157	156				
α -LP			T	I	T	G	C	Q	Y	G			Strand 2	
Elastase			V	T	P	L	Y	A	Q	Q				
Δ			1.38	0.42	0.76	2.68	1.08	1.00	1.35	1.20				
	180	181	182	183	184									
α -LP	L	T	Q	G	N								Strand 3	
Elastase	M	V	C	A	G									
Δ	1.32	1.33	1.01	1.49	1.21									
			(A)	(B)										
	191	192	192	192	193	194	195	196	197	198	199	200		
α -LP	C	M	G	R	G	D	S	G	G	S	W	I	Strand 4	
Elastase	S	G	C	Q	G	D	S	G	G	P	L	H		
Δ	1.41	0.99	1.75	1.06	0.87	0.29	0.25	0.82	1.22	0.45	0.26	0.79		
					216	215	214	213	212	211	210	209	208	207
α -LP					G	G	S	M	V	G	Q	A	Q	G
Elastase					V	F	S	T	V	G	H	V	A	Y
Δ					0.97	0.94	0.73	0.60	1.14	1.33	0.93	1.63	1.52	1.09
					225	226	227	228	229	230	231			
α -LP					S	S	L	F	E	R	L			
Elastase					P	T	V	F	T	R	V			
Δ					1.82	0.83	0.73	0.37	0.50	1.72	1.18			

Those residues enclosed in boxes are chemically identical in α -lytic protease and elastase. The amino acid sequences are expressed in the 1 letter code³⁰ and have been taken from ref. 20. Δ is the r.m.s. deviation of the α -carbon positions after optimisation and is expressed in Å units. The average r.m.s. deviation for domain I is 0.96 Å; for domain II, 1.12 Å. Those residues which are buried (internal) have a line drawn over the residue number. The strands are drawn diagrammatically to represent the antiparallel β hydrogen bonding seen in Fig. 1.

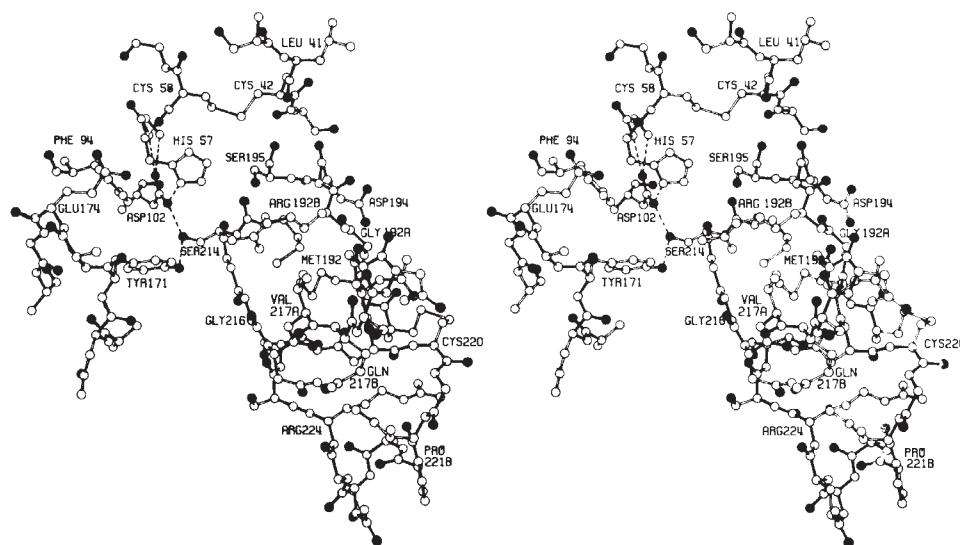


Fig. 2 Stereo drawing of the active site of α -lytic protease. The polypeptide main chain bonding is shown with solid black bonds. All oxygen atoms present are distinguished by solid black circles. Hydrogen bonds involving active site residues are illustrated as dashed lines. The side chains of Met 192, Met 213 and Val 217A occlude the primary binding pocket. The conformational lability of Arg 192B would not contribute to such a blockage of the substrate binding site.

side chains (Ala, Val, Ser) are quite different and were not predicted from model building¹². The active site and substrate binding region of α -lytic protease is shown in Fig. 2. The primary specificity site of the pancreatic-like serine enzymes involves the residues from 191–195 and 214–220. Elastase achieves its specificity because of the presence of Val 216 and Thr 226 (ref. 13), both side chains of which block the entrance to a hydrophobic cavity that exists in α -chymotrypsin⁴; the topologically equivalent residues in α -chymotrypsin are Gly 216 and Gly 226 (ref. 20). Residue 216 is also a glycine in α -lytic protease and therefore the structural reason for its specificity must differ from that of elastase and from that proposed¹².

Two important insertions in the sequence of α -lytic protease as compared with that of elastase occur in the active site region²⁰. The first of these insertions is a dipeptide at position 192. As a consequence, the side chain of Met 192 occludes the specificity pocket of α -lytic protease, whereas the side chain of Glu 192 in elastase is directed away from this binding pocket. The second insertion is a pentapeptide at position 217. The side chain of Val 217A of this insertion is also directed into the specificity pocket (see Fig. 2). Consequently, the primary binding site of α -lytic protease has room for substrate residues with side chains only as large as that of a valine residue¹⁶. In the region of residues 214–220, McLachlan and Shotton had misaligned (by four residues) the sequence of α -lytic protease with that of elastase and had placed a glutamine residue at position 216 and concluded that this residue in α -lytic protease would diminish the size of the primary specificity site, as does Val 216 in elastase. However, this glutamine residue is really residue 217B in the sequence and is located on the surface of the enzyme with its side chain pointing away from the active site (Fig. 2). Moreover, that misalignment had placed an asparagine residue at position 214 whereas a topologically equivalent serine residue has been found at this site in all of the serine protease structures that have been reported to date^{16–19,23,26,28} (including the evolutionarily unrelated subtilisin²⁹).

The extensive degree of structural similarity of these two serine proteases, in spite of the lack of sequence identity of topologically equivalent residues, makes it necessary to evaluate the expectations of a structural prediction. Our comparison of the 2.8-Å structure of α -lytic protease with the 2.5-Å structure of elastase shows that the molecules are very similar in the regions near those residues important for catalysis. However, it does not seem possible to rely on primary structural homology (or lack thereof) to predict correctly and accurately an unknown structure from that of an isofunctional protein molecule of known tertiary structure. Resort to secondary structure predicting rules^{1–5} would not necessarily offer much help, for in the stretch 131–138 of α -lytic protease an α -helical conformation is

incorrectly predicted whereas the corresponding region of elastase is indifferent to either α or β structure¹. If the results of protein structure prediction are to be used only to detect the possibility of conformational similarity and thus evolutionary homology, then such techniques can be useful^{12,15}. However, if the interest is in the structural detail of substrate binding sites, or, in the conformation of the catalytically important residues, then the present methods of predicting structure are not satisfactory.

This work was funded by a grant from the MRC of Canada to the Protein Structure and Function Group (to M.N.G.J.) at the University of Alberta. G.D.B. was supported by an MRC studentship.

LOUIS T. J. DELBAERE
GARY D. BRAYER
MICHAEL N. G. JAMES

Department of Biochemistry,
University of Alberta, Edmonton,
Alberta, Canada T6G 2H7

Received 30 October 1978; accepted 16 March 1979.

1. Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 222–245 (1974).
2. Lewis, P. N., Momany, F. A. & Sheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2293–2298 (1971).
3. Lim, V. I. *J. molec. Biol.* **88**, 873–878 (1974).
4. Finkelstein, A. V. & Ptisyn, O. B. *J. molec. Biol.* **62**, 613–616 (1971).
5. Kabat, E. A. & Wu, T. T. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1473–1478 (1973).
6. Schulz, G. E. *et al. Nature* **250**, 140–142 (1974).
7. Matthews, B. W. *Biochim. biophys. Acta* **405**, 442–451 (1975).
8. Schulz, G. E. *Angew. Chemie (Int'l Edn)* **16**, 23–32 (1977).
9. Levitt, M. & Warshel, A. *Nature* **253**, 694–698 (1975).
10. Kuntz, I. D. *et al. J. molec. Biol.* **106**, 983–994 (1976).
11. Ptisyn, O. B. & Rashin, A. A. *Biophys. Chem.* **3**, 1–20 (1975).
12. McLachlan, A. D. & Shotton, D. M. *Nature* **229**, 202–205 (1971).
13. Shotton, D. M. & Watson, H. C. *Nature* **225**, 811–816 (1970).
14. Birktoft, J. J., Blow, D. M., Henderson, R. & Steitz, T. A. *Phil. Trans. R. Soc. B* **257**, 67–76 (1970).
15. Kretsinger, R. H. & Barry, C. D. *Biochim. biophys. Acta* **405**, 40–52 (1975).
16. Brayer, G. D., Delbaere, L. T. J. & James, M. N. G. *J. molec. Biol.* (in the press).
17. Delbaere, L. T. J., Hutcheon, W. L. B., James, M. N. G. & Thiessen, W. E. *Nature* **257**, 758–763 (1975).
18. Brayer, G. D., Delbaere, L. T. J. & James, M. N. G. *J. molec. Biol.* **124**, 261–283 (1978).
19. Delbaere, L. T. J., Brayer, G. D. & James, M. N. G. *Can. J. Biochem.* **57**, 135–144 (1979).
20. James, M. N. G., Delbaere, L. T. J. & Brayer, G. D. *Can. J. Biochem.* **56**, 396–402 (1978).
21. Olson, M. O. J., Nagabhushan, N., Dzwiniel, M., Smillie, L. B. & Whitaker, D. R. *Nature* **228**, 438–442 (1970).
22. Kaplan, H. & Whitaker, D. R. *Can. J. Biochem.* **47**, 305–316 (1969).
23. Sawyer, L. *et al. J. molec. Biol.* **118**, 137–208 (1978).
24. Rossmann, M. G. & Argos, P. *J. biol. Chem.* **250**, 7525–7532 (1975).
25. Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T. & Xuong, N. H. *Biochemistry* **9**, 1997–2009 (1970).
26. Bode, W., Schwager, P. & Huber, R. *J. molec. Biol.* **118**, 99–112 (1978).
27. Narayanan, A. S. & Anwar, R. A. *Biochem. J.* **114**, 11–17 (1969).
28. Birktoft, J. J. & Blow, D. M. *J. molec. Biol.* **68**, 187–240 (1972).
29. Matthews, D. A., Alden, R. A., Birktoft, J. J., Freer, S. T. & Kraut, J. *J. biol. Chem.* **252**, 8875–8883 (1977).
30. *Biochem J.* **113**, 1–4 (1969).

matters arising

Human activity and the erosion of soils on chalk

In their letter¹ relating to beech bark disease and lynchets, Lonsdale *et al.* refer to the probability of an acid non-calcareous soil overlying the chalk on their site. This type of soil seems to be very widespread in the west Sussex-east Hampshire region, more so than the chalky soil which is generally supposed to be the characteristic one, and is by no means always underlain by Pleistocene or earlier deposits. In fact, on shallow slopes, a highly calcareous soil seems to occur only where human disturbance has facilitated erosion; this usually means in ploughed fields or on the negative aspects of lynchets such as Lonsdale *et al.* have observed. Steep slopes, including the chalk scarp, are, on the other hand, often calcareous. The wood in which Lonsdale *et al.* are working does indeed seem to contain the non-calcareous soil, although not in its most acid form; samples taken from it² in 1970, outside the lynchet area, were mostly found on analysis to have a non-calcareous, non-iron, mineral content of ~70%. They are significantly less calcareous than samples from the neighbouring ploughed fields, and this difference has been found in several such areas of old woodland cover and arable, even though the two resemble each other in topographic and other respects.

M. A. COLLINS

Geography Department,
London University
King's College Field Centre,
Rogate, nr Petersfield,
Hampshire, UK

1. Lonsdale, D., Pratt, J. E., & Aldsworth, F. G. *Nature* **277**, 414 (1979).

2. Collins, M. A. *History and Soils on the South Downs Rogate Papers* no. 1 (1978).

Aerosol anomalies preceding earthquakes

In his article on aerosol anomalies preceding earthquakes Tributsch¹ suggests that there is emission of positive ions

before an earthquake occurs and that these ions are responsible for the many reports of animal disturbance before the major event itself. We have been investigating electrical phenomena associated with stress and failure in rocks and ceramics and some of our data may be applicable to Tributsch's hypothesis.

For testing, we normally attach connections to the rock via conductive epoxy or hammered-in phonograph needles. These points are connected to one or more electrometers before the specimen is placed in compression or bending. As the load is applied we see a series of pulsed currents that we associate with piezoelectric effects. As the load is increased there is a steady state current of the order of 10^{-10} A that we associate with the stress induced migration of ions² or the inception of a continuous series of internal microfractures. In any case the current pattern and areas of highest current can be used to predict areas of high residual stress and the location of ultimate failure.

As the specimen approaches failure we observe the emission of both electrons and positive ions with energies up to 15 eV at current levels of 10^{-11} A. If the experiment is done in a darkened room there are flashes of light as the material begins to fail. These flashes are not observed in a helium environment but are intensified in a water vapour or pure oxygen environment; we suggest that the electrons and ions are exciting water or oxygen molecules that decay by photon emission. This electron and ion emission may well be responsible for the effects discussed by Tributsch as well as the earthquake light phenomena reported by Derr³.

A full report on this work will be given at the Fourth International Congress on Rock Mechanics in September at Montreux.

This work was supported by the Engineering Division of the NSF.

STUART A. HOENIG

Department of Electrical Engineering,
The University of Arizona,
Tucson, Arizona 85721

1. Tributsch, H. *Nature* **276**, 606 (1978).

2. Weber, N. & Goldstein, N. *J. chem. Phys.* **41**, 2898–2901 (1964).

3. Derr, J. *Bull. seismol. Soc. Am.* **63**, 2177–2187 (1973).

Historical climatology

WE should like to join Professor H. H. Lamb in clarifying a point in our article on historical climatology¹. We stated that the correlation between Lamb's English winter severity index values for the decades between AD 1100 and 1400 and Alexandre's decadal indices for winters in Belgium was very low. We attributed this largely to the fact that Alexandre used thoroughly criticised and carefully evaluated historical data, whereas the sources available to Lamb when he pioneered the method were compilations of variable but generally poor quality.

It is necessary to add that in using the indices to reconstruct temperatures, Lamb himself expressly avoided undue emphasis on the decadal values as such, instead relying on half-century mean values derived from them². These smoothed values correlate much better with Alexandre's series. It is important to emphasise Professor Lamb's caution in this respect. Indeed, it was because his caveat evidently required reiteration that we originally chose to highlight the problems inherent in the decadal values.

M. J. INGRAM

D. J. UNDERHILL

T. M. L. WIGLEY

H. H. LAMB

Climatic Research Unit,
University of East Anglia,
Norwich, UK

1. Ingram, M. J., Underhill, D. J. & Wigley, T. M. L., *Nature* **276**, 329–34 (1978).

2. Lamb, H. H., *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **1**, 13 (1965).

The impossibility of comminuting small particles by compression

KENDALL¹ has given a simple explanation of the impossibility of comminuting and crushing a brittle body by cracking when its size is smaller than a certain critical value. This size effect was demonstrated by applying the classical Griffith²

energy criterion to a brittle body of idealised geometry. However, it has been shown recently³ that the Kendall theory could be improved, and brought into complete agreement with the Griffith theory, if allowance were made for the restraint to the free lateral movement inevitably experienced by the brittle body during compression. Assuming the lateral restraining force to be a small fraction, α , of the axial compression, F , and following exactly Kendall's¹ procedure and notation, the cracking force is

$$F = \frac{1}{(1 - w/d + 8\alpha c/d)} \left(\frac{2ERd}{3} \right)^{1/2} \quad (1)$$

where c is the crack length.

The force at the transition from cracking of the brittle body to its gross yielding under the platen is given by the solution to the quadratic

$$\frac{1}{Yd} \left(\frac{F}{b} \right)^2 - \frac{F}{b} (1 + 8\alpha c/d) + \left(\frac{2ERd}{3} \right)^{1/2} = 0 \quad (2)$$

In particular, if the particle size is reduced to a critical value

$$d_{crit} = \frac{32ER}{3Y^2(1 + 8\alpha c/d)^4} \quad (3)$$

cracking becomes impossible. Equations (1)–(3) differ from their counterparts in Kendall's paper by the presence of the very crucial term $8\alpha c/d$.

Based on my experimental data³ and that of Kendall⁴, I have argued that the restraining force would be very small especially if the size of the body is reduced. In fact, a value of α between 1/100 and 1/200 seemed most appropriate.

For the model material (polystyrene) tested by Kendall¹ the c/d ratio in equation (3) for all samples was 1.25. Equation (3) suggests that for $\alpha = 1/200$ the transition from brittle to ductile behaviour of polystyrene could be expected at a size of 3.69 mm as opposed to 4.48 mm predicted by Kendall's theory. The experimentally observed value was 3.6 mm (ref. 1). Expression (3) clearly is in better agreement with experimental data. Likewise, equation (3) predicts a crushing limit of around 0.85 μ m for calcium carbonate which is also in close agreement with the average particle size of 0.8 μ m to which calcium carbonate is reduced after prolonged milling.¹

B. L. KARIHALOO

Department of Civil Engineering,
University of Newcastle,
New South Wales 2308, Australia

Agonist regulation of α -adrenergic receptor numbers

THERE have recently been several reports which indicate that agonist-induced desensitisation of cells might be explained, at least in part, by reductions in the number of receptors exposed at the surface of target cells; the evidence has been particularly clear for β -adrenergic receptors. Thus, it is not surprising that studies have been designed to see whether agonist-induced desensitisation of cells to α -adrenergic stimulation might be explained in a similar way. Two reports have now described fairly rapid agonist-induced decreases in α -receptor number^{1,2}, and another described a modest rise in α -receptor number in animals treated for several weeks with reserpine to induce supersensitivity³.

In their report, Cooper *et al.*² seemed to consider that the observed reduction in receptor number provided an adequate explanation for the desensitising effects of agonists; no other mechanisms were apparently considered. However, this position is untenable, as their main observation was that a 50% decrease in receptor number in adrenaline-treated platelets (assessed by dihydroergocryptine binding) was accompanied by an essentially complete loss of two cellular responses to the same agonist (aggregation and 5-hydroxytryptamine secretion).

Surely, this must mean that only a small proportion of the desensitising effect of agonist treatment can have been mediated through the change in receptor number? Most of this desensitising effect was presumably a result of some event which was brought about by receptor activation but which did not involve any changes in the number of exposed receptors. This is hardly surprising, as there is much evidence which indicates that in other tissues, especially smooth muscles, changes in receptor numbers are not likely to be the main mechanism whereby increased or decreased exposure to agonists brings about changes in cell sensitivity to those ligands which act at Ca^{2+} -mobilising receptors^{4–6} (of which the α -adrenergic receptor is one type⁷). These alternative mechanisms for changes in sensitivity may not be as conceptually simple nor as well understood as the changes in receptor number, but they cannot be ignored when the available data are incompatible with an explanation based solely on changes in receptor number.

ROBERT H. MICHELL

Department of Biochemistry,
University of Birmingham,
PO Box 363, Birmingham, UK

- U'Pritchard, D. & Snyder, S. H. *Eur. J. Pharmac.* **51**, 145–155 (1978).
- Fleming, W. W., McPhillips, J. J. & Westfall, D. P. *Ergebn. Physiol.* **68**, 55–119 (1973).
- Carrier, O. *Fedn Proc.* **34**, 1975–1980 (1975).
- Fleming, W. W. *Rev. Neurosci.* **2**, 43–90 (1976).
- Jones, L. M. & Michell, R. H. *Biochem. Soc. Trans.* **6**, 673–688 (1968).

ALEXANDER AND HANDIN REPLY—Michell is correct, of course, in stating that agonist-induced desensitisation of physiological responses might not be explainable on the basis of a 40–50% decrease in receptor number, as determined by ligand binding with ³H-dihydroergocryptine (DHEC), an α -adrenergic antagonist. It was not our intention to imply that this was the sole explanation for adrenaline-induced desensitisation in the platelet. However, we do not agree “that only a small proportion of the desensitising effect of agonist treatment can have been mediated through the change in receptor number”. Since submission of our paper, there has been a report on the desensitisation of β -adrenergic receptors showing that the percentage of receptors identified by agonist binding is decreased to an extent considerably greater than the 35% decrease in receptor number determined by antagonist binding. These data suggest that agonists and antagonists bind different forms of the β -adrenergic receptor¹. It has been reported that α -adrenergic agonists and antagonists bind to distinct states of the α -adrenergic receptor and that ³H-DHEC may label both the agonist and antagonist state^{2–5}. If most of the decrease in ³H-DHEC binding in platelets preincubated with (–)adrenaline represents changes in the putative agonist state, then the observed decrease in receptor number may be sufficient to explain most of the α -adrenergic agonist-induced desensitisation in platelets. This possibility requires further investigation using both agonist and antagonist binding assays.

The references on denervation supersensitivity may not be relevant to our work on desensitisation. In fact, it has recently been reported that α -adrenergic receptors in rat cerebral cortex increase after chronic reserpine treatment, suggesting that an increase in receptor number may contribute to post-junctional supersensitivity⁶.

R. WAYNE ALEXANDER
ROBERT I. HANDIN

Peter Bent Brigham Hospital,
Boston, Massachusetts 02115

- Wessels, M. R., Mullikin, D. & Lefkowitz, R. J. *J. biol. Chem.* **253**, 3371–3373 (1978).
- Greenberg, D. A. & Snyder, S. H. *Molec. Pharmac.* **14**, 38–49 (1978).
- Peroutka, S. J., Greenberg, D. A., U'Pritchard, D. C. & Snyder, S. H. *Molec. Pharmac.* **14**, 403–412 (1978).
- U'Pritchard, D. C., Charness, M. E., Robertson, D. & Snyder, S. H. *Eur. J. Pharmac.* **50**, 87–89 (1978).
- Miach, P. J., Dausse, J. P. & Meyer, P. *Nature* **274**, 494–497 (1978).
- U'Pritchard, D. C. & Snyder, S. H. *Eur. J. Pharmac.* **51**, 145–155 (1978).

- Kendall, K. *Nature* **272**, 710–711 (1978).
- Griffith, A. A. *Phil. Trans. R. Soc. A* **221**, 163 (1920).
- Karihaloo, B. L. *Proc. R. Soc. (in press)*.
- Kendall, K. *Proc. R. Soc. A* **361**, 245–263 (1978).

- Strittmatter, W. J., Davis, J. N. & Lefkowitz, R. J. *J. biol. Chem.* **252**, 5478–5482 (1977).
- Cooper, B., Handin, R. I., Young, L. H. & Alexander, R. W. *Nature* **274**, 703–706 (1978).

reviews

Concepts of colour

J. D. Mollon

Color Vision: An Historical Introduction. By G. S. Wasserman. Pp. 224. (Wiley: New York and Chichester, UK, 1978.) £13.40.

THOSE who study colour vision are very aware of the history of their field. There are two reasons for this. First, the topic is one that has recommended itself to the strongest intellects: Newton, Thomas Young, Maxwell and Helmholtz have left writings that set models of experimental elegance and theoretical clarity for the modern researcher. Second, the advancement of the field has depended on an instructive sequence of conceptual insights, which must be recapitulated by each fresh student. It is therefore very appropriate that Professor Wasserman should adopt an historical organisation in his intermediate-level textbook on colour vision.

As actual history, Wasserman's account is patchy and largely limited to expositions of the main printed sources. His summary of the relevant parts of Newton's *Opticks* is accurate and very clear; but his treatment of other early material is unashamedly based on secondary sources and he entirely omits mention of the trichromatists of the first half of the eighteenth century. His account of the contribution of Maxwell and Helmholtz is excellent, but then he spoils himself by recommending that the three-receptor theory of colour vision, which is widely referred to as the Young-Helmholtz theory (and which is almost certainly correct), should instead be eponymously linked with Newton and Maxwell. That Maxwell should enjoy more credit is fair; but to associate the theory with Newton is to confound the whole history of the field. The most fundamental fact about colour vision is that of trichromacy, the fact that only three variables are needed in a colour-matching experiment; and it is sometimes said that trichromacy is implicit in Newton's colour circle and centre-of-gravity rule. But (as Wasserman himself very clearly explains) by mixing three primary colours that lie

on the circumference of Newton's circle, one can reach only that part of colour space enclosed by a triangle inscribed within the circle. Newton lacked the concept of imaginary primaries lying outside the colour circle; and he lacked that concept because he lacked in turn the idea of narrowly tuned transducers, supposing only that "... the Rays of Light in falling upon the bottom of the Eye excite vibrations in the *Tunica Retina*" (Qu. 12) and that these vibrations are "propagated through the solid, pellucid and uniform Capillamenta of the optic nerves into the place of sensation" (Qu. 23). The "several sorts of Rays" merely "make Vibrations of several bignesses, which according to their bigness excite Sensations of several Colours" (Qu. 13). As Newton understood neither the fact of trichromacy nor the trichromatic theory that was later introduced to explain it, I urge colleagues not to adopt Wasserman's suggestion that we should speak of the Newton-Maxwell theory.

The topics covered by Wasserman include colorimetry, colour blindness, component and opponent theories, zone theories, photometry, physiology. There are detailed discussions of the theories of Hurvich and Jameson and of Guth; of Land's experiments on colour contrast; of heterochromatic photometry; and of the early microspectrophotometric records. Wasserman has taken great pains to elucidate conceptual difficulties for students and this trouble is repaid by a very lucid text. I very much liked his discussion of imaginary primaries and his account of non-additivities of heterochromatic lights. There is a very helpful passage relating Granit's early reports of retinal 'dominators' and 'modulators' to the later electrophysiology.

But the book has two weaknesses. First, the coverage is unbalanced (the British come off especially badly). Rushton's reflection densitometry is dismissed as valueless in a sentence. Stiles' increment-threshold method passes entirely unmentioned, although the spectral sensitivities so derived are

close to linear transformations of the colour-matching functions and (from this evidence and from microspectrophotometry) now seem likely to represent the sensitivities of the cones themselves. (Indeed, although the book is illustrated with a great number of spectral sensitivities, amazingly there is nowhere shown a set that resembles the ones likely to be correct.) Nothing is said of one of the most interesting developments of the last decade, Zeki's demonstration of areas in the prestriate cortex that seem to be specialised for the analysis of colour.

Second, Wasserman has some heresies of his own to press upon the unsuspecting student. Thus the chapter on colour blindness is quite out-of-date and misleading and the author perpetuates one of the most notorious conceptual errors of the field, that of supposing that anomalous trichromacy can be explained by giving abnormal weightings to the signals from the three cones of normal vision. There is a lot of dubious talk about two-peaked photo-pigment sensitivities in invertebrates and in vertebrates, without any mention to the student that invertebrate pigments may be bistable (the two states corresponding superficially to a rhodopsin and a metarhodopsin) or that the evidence for visually significant short-wave peaks in vertebrate pigments is very slight. The recent microspectrophotometric measurements of goldfish receptors shown on p189 do indeed exhibit secondary *cis*-peaks, but the goldfish pigments are based on vitamin A₂ and Wasserman does not point out that the *cis*-peak of the long-wave human pigment (P565) would lie well below 400 nm, at wavelengths completely absorbed by the lens of the eye.

In summary, Wasserman's book is good in parts. Should this Curate's Egg be given to students? Only if accompanied by many, well placed pinches of salt. □

J. D. Mollon is Lecturer in Experimental Psychology at the University of Cambridge, UK.

Biological properties of immunoglobulins

Comprehensive Immunology. Volume 5: Immunoglobulins. Edited by R. A. Good and G. W. Litman. Pp. 381. (Plenum Medical: New York and London, 1978.) \$18.58.

THIS is a multi-authored book containing 16 different articles. The style of each is governed more by the authors themselves than by any unifying editorial criterion. The length of the chapters is also very variable and bears no relationship to their general importance. Whether this is by chance or design, it turns out to be fortunate because, in general, the most extensively, and to my mind the better, reviewed subjects have also suffered the least from the unavoidable delay in publication of books of this nature.

The major emphasis is on the more classical immunochemistry problems concerned with the nature of the antibody-antigen interaction. Crystallographic, physicochemical, kinetic and thermodynamic data bring us a comprehensive picture of the general properties of the antibody-combining sites and, more importantly, of the outstanding and unresolved aspects of the problem. For those of us trying to understand the peculiarities of monoclonal antibodies in serological reactions, some of the chapters (particularly that by Karush on "The Affinity of Antibody: Range, Variability, and the Role of Multivalence") are very relevant. Yes, Cathou was right in saying, "Immunochemistry does not appear to be in any danger of an early death" (p76).

The first six chapters, which also include discussions on the structural basis of the effector functions of antibodies and other biological properties, take up well over half of the book. The other portion doesn't attempt a comprehensive coverage of other aspects of immunoglobulins, but a quick glance at the table of contents may give a misleading impression of the real subject matter. For instance, the short discussion by Ohno on the significance of gene duplication in Ig evolution is intended for the initiated and not to give a picture of the complexities of the contraction and expansion processes in the evolution of immunoglobulin gene pools. On the other hand, I found a considerable amount of information on subjects that are seldom reviewed: for example, on immunoglobulin-like molecules and possible functional ancestors of antibodies in lower species (chapter 8) and on the relative expression of V_H subgroups in

normal and abnormal immunoglobulins in man (chapter 13).

Chapter 10 reviews the structure of atypical immunoglobulins and their implications in terms of genetic control mechanisms. It is a great pity that Frangione did not have an extra year to write this chapter, as his pioneering ideas about a separate 'gene' for the hinge region could have crystallised with the discoveries of 'gaps' in the DNA structure. The book closes with a crisp "Overview on membrane immunoglobulins": the outstanding problems, the relevant facts and the prevailing opinions at the time of publication. Pernis has obviously made a great effort to try to include last-minute references and, in the rush, Berget *et al.* (1977) did not make it to the reference list.

A number of errors have crept into

Field theory

Relativistic Quantum Fields. By C. Nash. Pp.223. (Academic: London, New York and San Francisco, 1978.) £15; \$31.

IN the past five years we have witnessed an astonishing renaissance of field theory in elementary particle physics. On the one hand the Weinberg-Salam unified theory of weak and electromagnetic interactions has had considerable success in fitting a large variety of neutrino neutral current data; on the other hand quantum chromodynamics (QCD) predicted and explains the scaling violations observed in deep inelastic neutrino scattering. As a result it is now the conventional wisdom that these theories, or something quite like them, will eventually permit a full understanding of all of the fundamental interactions seen in nature, except perhaps gravity.

Such an understanding is still a long way off and it will come, if at all, only when our facility in using and thinking within the field theoretic formalism has advanced considerably beyond the present level. (I am thinking, for example, of the problem of quark confinement—plainly a non-perturbative effect which will need considerable technical progress to elucidate.) It is therefore a pleasure to welcome an excellent book by Charles Nash which develops some of the fundamentals of field theory with enough detail to teach the serious student how to undertake his or her own calculations. Most of the book is devoted to the $\lambda\phi^4$ theory and quantum electrodynamics (QED), rather

than the non-abelian gauge theories (Weinberg-Salam and QCD) which launched the revolution. This is all to the good, as it permits the assimilation of unfamiliar techniques without the encumbrances of the new theories.

The first *sine qua non* of the new order is functional differentiation and integration, which is illustrated by the usual Gaussian integral. (Is this the *only* functional integral anyone has done?) Nash takes us gently through this, and even consoles us should we feel "a little perplexed or confused". "This is quite normal". Functionals on a Grassmann algebra are also considered. Next there is an extensive coverage of dimensional regularisation, a covariant and gauge invariant method for handling the infinite quantities which inevitably arise when evaluating radiative corrections in these theories. This is used to verify the Ward Identity and to compute the renormalisation constants in QED. The final chapter touches on the asymptotically free gauge theories, the Callan-Symanzik equation, and so on. However, the coverage is less full than that of the earlier material, and there is no treatment of spontaneous symmetry breaking in non-abelian theories. If these attractive theories do withstand further rigorous experimental tests, it is to be hoped that the author will be moved to write a companion volume dealing with these topics in the same detail. In the meantime the present book will certainly provide a lastingly useful and relatively painless introduction into a highly technical field.

César Milstein

César Milstein is Head of the Protein Chemistry Subdivision, Medical Research Council Laboratory of Molecular Biology, Cambridge, UK.

D. Bailin

D. Bailin is Lecturer in Theoretical Physics at the University of Sussex, Brighton, UK.

Amino acid analysis

Amino Acid Determination: Methods and Techniques. Second edition. Edited by S. Blackburn. Pp. 367. (Marcell Dekker: New York and Basel, 1978.) Swiss Fr. 82.

THIS will be a most useful book which should find its way on to the shelves of every amino acid analyst. Old hands will find new and valuable information, and beginners will find a thorough introduction to the field. Perhaps its best features are the extensive and thorough bibliography which accompanies each chapter and the author index, covering over 1,000 authors. Nevertheless, and despite its subtitle, this is *not* a laboratory manual in the sense of providing recipes and precise details of procedural matters.

There are 13 chapters of which the editor has written a short introduction and recent history (chapter 1), an equally short tailpiece (chapter 13) and three other chapters. Of the remaining eight chapters, five are by George W. Robinson and the others by experts in their particular areas.

Chapter 2 (by Blackburn) describes the preparation and hydrolysis of samples before analysis and is a useful survey of the methods and pitfalls in this necessary preliminary. There follow four chapters (3–6) by Robinson on the standard ion-exchange methods used in modern automatic analysis, dealing with the theory, development, and practical aspects such as resins, buffers and detection systems. Chapter 7, entitled "Automatic Analysers", is the least satisfactory in the book. Admittedly, information on makes of instruments and their capabilities and defects is necessarily ephemeral material but it is nevertheless very valuable information for anyone undertaking the purchase of a new analyser. This chapter could therefore have said much more about the strengths and weaknesses of the different instruments and their makers. For example, although the editor mentions the Rank-Hilger Chromospek in chapter 1 as possessing some "novel features", neither the instrument nor its features are mentioned in chapter 7. In fact, novel though they may be, there are serious disadvantages that could well have been discussed and compared with the pros

and cons of other makers' instruments.

Blackburn's chapter 8 on "Gas Chromatography" will be found useful by GC buffs no doubt, but for a method which, in the words of the Editor in chapter 1, "has failed to displace the amino acid analyser" this chapter is extraordinarily long (78 pages compared to a total of only 69 in chapters 3–7 on the analyser). This may be necessary to do the method justice, as it does have specialist uses; but the balance seems wrong.

The use of computers in amino acid analysis is well covered in chapter 9, and chapter 10 is a rag-bag of miscellaneous procedures and applications containing a useful section on dansyl amino acids. An unexpected blessing, at least for those with medical interests, is chapter 11 on "Physiological Fluids". This is a real mine of densely packed information. Finally, chapter 12 is a useful account of the relatively new techniques using pyridoxal derivatives of amino acids.

W. Ferdinand

W. Ferdinand is Senior Lecturer in Biochemistry at the University of Sheffield, UK.

Viking missions

To the Red Planet. By Eric Burgess. (Columbia University Press: New York, 1978.) \$24.95.

ALTHOUGH all the plates and other illustrations in this book are in black and white, this is more than compensated for by the actual choice of graphic material and the informed, balanced and stimulating text. The opening chapters give a lucid account of the historical background and current state of knowledge regarding the planet Mars, the biophysical and biochemical bases of life as we know it, and how these topics relate to the objectives of the Viking missions.

Passages devoted to spacecraft instrumentation and hardware, which appear mainly in chapter 4, are treated in the contexts of the scientific rationale and mission planning. The emphasis here, as throughout the book, is sufficiently human in orientation that one does not lose sight of the involvement, the thoughts and reactions, of the project scientists; and the author pays due regard to maintaining the requisite attention to scientific and technical detail.

Burgess arranges his facts well and combines them with lively narrative and wit, not born of offhandedness nor

any trace of cynicism, but of the humour flexed by an author who really is in control of his material and who obviously feels a pressing need to communicate this to a wide spectrum of readers.

Each point of discussion is given a fresh paragraph, assisting the overall clarity, but a few sub-headings would not have gone amiss, particularly where the Viking lander experiments are described in chapter 7 and where conclusions are drawn in the final chapter. But this is not a serious detraction. Possibly through modesty the author may not have regarded his work as a book of reference. In view of the selection of topics and pictures discussed, and the high standard of his own exposition, it certainly could be regarded as such.

Burgess has undertaken a difficult task and performed it well; readers of the book will undoubtedly be rewarded. If they happen to be members of the public at large or specialists in planetary science, it will go a long way towards broadening their appreciation of the considerable amount of interdisciplinary thought, effort and coordination which has contributed to the success of the Mars Viking missions.

D. J. Telfer

D. J. Telfer is a Research Associate in the Lunar and Planetary Unit, Department of Environmental Sciences, University of Lancaster, UK.

SCIENTIFIC BOOKSHOP

H. K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.50. (Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

**H. K. LEWIS
& Co. Ltd.**

136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavit,
London, WC1E 6BS."

Circle No. 04 on Reader Enquiry Card.

Ageing of higher organisms

Principles of Mammalian Ageing. Second edition. By R. R. Kohn. Pp. 240. (Prentice-Hall: Englewood Cliffs, New Jersey, and Hemel Hempstead, UK, 1978.) \$19; £13.85.

GERONTOLOGY was born in optimism. It threatens to founder in verbiage. Papers and books are now appearing in quantity, while declining in quality, drowning the few really important studies in this field. Robert Kohn's book was there at the beginning of the boom, and it stands aside from the information explosion. Despite certain problems, this new edition restores to us a compact, well-written and refreshingly intelligent introductory text on the ageing of higher organisms.

The book begins by exploring the progressive crystallisation and toughening of inorganic materials as an analogy to illustrate ageing. It is an analogy which some may find weak and misleading, for ageing is manifest by increasing disorder rather than order. Nevertheless it serves the author's purpose, for he moves on to an excellent summary of the ageing phenomena in the extracellular materials of the body. The following sections deal with ageing within cells, briefly covering the evidence for age-related changes in the dynamic renewal of cellular constituents and for long-term biochemical deterioration. This is followed by a carefully ordered chapter on ageing of cell populations, which classifies cells according to their degree of continuing cell division within tissues. The rest of the book is concerned with the ageing of animals and here the author makes a good job out of the onerous task of summarising a highly complex and contradictory field.

The problem with the book is the author's penchant for selecting two or three seemingly contradictory papers as representative of major fields of ageing research, which then allows him to conclude that this work is inconclusive. For example, there is excellent evidence that dividing cell populations such as mammary gland cells, antibody-producing cells or haematopoietic cells show some age changes during the lifespan of the individual and will eventually cease division, although they are capable of outliving the organism from which they are derived. The author fails to mention significant work in this field and uses sparse studies to support a 'no-evidence' conclusion. He does the same with the exciting new field of the analysis of errors in macromolecules of senescent cells, with work on histones, on the question of protein

turnover, repair capacity, DNA synthesis and the genetics of ageing as a whole. A more penetrating analysis would have shown him that work which at first reading seems contradictory is in fact complementary, and would have enabled him to pick out the key papers in each field. This would have allowed him to draw constructive conclusions in each area, thus enriching the book.

The author's unwillingness to make conclusions derives in part from a certain failure to properly grasp molecular, genetic and evolutionary concepts as a whole. He remakes the classic *faux pas* that ageing exists because "old, non-reproductive animals compete with younger animals for food", which is the same as saying that ageing exists because of ageing. Besides this, scientists usually have some inner vision which guides their work, especially in such a complex and

exploratory field as ageing research. To some, ageing is a process of headlong over-differentiation. To others it is an ineluctable entropic deterioration of biological information. Krohn seems to see ageing as a kind of crystallisation or fossilisation of biological materials. Each man to his own. But it is important to be able to understand other researchers' ways of looking at the problem.

Yet it is easy to criticise a scientific text on the basis of selectivity. These criticisms do not alter the fact that the book holds together very well. Now that the number of papers on ageing is somewhere at the five figure level, a coherent and thoughtful introductory text such as this is no mean feat.

Stephen J. Fulder

Stephen J. Fulder is about to take up a research post at the Brookdale Institute of Gerontology, Jerusalem, Israel.

Transfer RNAs

Transfer RNA. Edited by S. Altman. Pp. 356. (MIT Press: Cambridge, Massachusetts, and London, 1979.) \$30; £21.

THE concept of tRNAs as the essential links between nucleic acid and protein, has for over twenty years been a central theme of molecular biology in general and studies on protein biosynthesis in particular. It is now some thirteen years since the 'state of the art' was comprehensively reviewed in a Cold Spring Harbor Symposium volume. Subsequent years have witnessed a veritable explosion in the literature, not only in classical areas of tRNA research but also in new ones such as analysis of tRNA genes, tRNAs as primers of reverse transcriptase activity, and determination of tRNA tertiary structure.

It is thus more than timely for an overview of all these aspects of tRNA function. This book succeeds admirably in this task, by providing a number of reviews, all by leading authorities in the field. The introduction (by Zachau) serves to establish the subject in a historical perspective, as well as reminding us of the many questions still left unanswered. The impact of our knowledge of the tertiary structure of just one tRNA is well illustrated throughout the book, and not only by the chapter on its crystal structure and immediate implications (by Kim). In particular, Clark's chapter collating features of primary, secondary and tertiary structure provides a fascinating account of the generalisations and clues to function at the molecular level provided by the yeast phenylalanine tRNA crystal

structure. The essay by Crothers and Cole on conformational changes of tRNA in solution serves to provide a physically based viewpoint complementary to descriptions of structure in the crystalline state.

The contribution of Iglio and Cramer on tRNA synthetases and their substrate interactions, elegantly describes the current status in this area, to which Cramer's group have themselves made many significant contributions. The beginnings of understanding at the molecular level are also discussed by Pongs in a chapter largely concerned with ribosomal and mRNA recognition in relation to protein synthesis. Both these chapters clearly indicate the important directions to pursue in the future, difficult though they will be.

Other sections in this book are more concerned with biological topics. Thus, Altman himself reviews the area of tRNA biosynthesis, and with Körner and Feinstein he discusses tRNA-mediated suppression. Many other roles of tRNA are reviewed by LaRossa and Söll, Nishimura, in a chapter on modified nucleosides and isoaccepting tRNAs, suggests that many of these regulatory roles involve modified residues, in as yet unrevealed ways.

Throughout this book the sense of excitement still pervading many areas of transfer RNA study, come across very clearly. It can thus be recommended as a worthwhile and stimulating read for graduate and even advanced undergraduate students in molecular biology as well as for those workers more actively involved in the subject.

Stephen Neidle

Stephen Neidle works in the Department of Biophysics, King's College, University of London, UK.